Activating Children's Thinking Skills (ACTS)

An Intervention Evaluation Study

Jessica Lucy Dewey

D.Ed.Psy. Thesis (Volume 1)

2006

University College London

UMI Number: U592021

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI U592021
Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.
All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Table of Contents

Abstract

This study investigated the impact of an infusion intervention on the cognitive, emotional and social development of Year 4 and Year 5 children across 10 schools in a large county. A two year intervention period was utilised to ascertain effects over time and contexts to monitor transfer and maintenance (Blagg, 1991; McGuinness, 1993; McKinstery & Topping, 2003). The project provided a unique contribution to the specialist sphere of thinking skills by using an *experimental* and matched *waiting list control* group to overcome limitations of earlier research design (Blagg, 1991; DfEE, 1999; Gorodetsky et al., 2002; Sternberg & Bhana, 1986). A multi-method pre, post and delayed post test using standardised measures across a range of learner outcomes developed research further in the domain of student effects in the primary age range (Coles, 1999; Nisbet, 1993; McKinstery & Topping, 2003; Watkins et al., 2001; Wilson, 2000). Furthermore, the research offered a distinctive insight into the perceptions of the key players involved through triangulated qualitative analysis which, unlike earlier studies examined the practicalities of the program in real life educational environments (Higgins et al., 2004; McGuinness, 2003; TES, 2002; Topping, 2002; Wilson, 2000).

Key findings demonstrate the impact of the ACTS intervention on enhancing children's cognitive abilities over a two year as opposed to a one year intervention period. This resonates with similar studies in which such an extended time period is considered essential to generate cognitive change (Blagg, 1991; Feuerstein et al., 1980; Shayer & Adey, 1993). Correspondingly qualitative data from students and staff illustrate that

children were not only able to detail the range of thinking skills taught but describe application. This evidenced transfer in both the near and far contexts an area ignored in prior work (Blagg, 1991; DfEE, 1999; McGuinness, 2003; Wilson, 2000). Additional effects linked the teaching of thinking to social and emotional gains of children in line with the evidence base (Blagg, 1991; Lipman et al., 1980; McKinstery & Topping, 2003). Considerable impact was, likewise, apparent in relation to the professional development of staff in terms of both skills and knowledge as paralleled in other studies (McGuinness et al., 1997; McKinstery & Topping, 2003; Munro, 1999; Stewart & Smardon, 2002).

The research is of significance in strengthening both the empirical and theoretical evidence base in the teaching of thinking skills by evaluating infusion. Important implications for both the psychological knowledge base and educational psychology practice are made. Of particular significance are the application of such programmes within curriculum contexts alongside the professional artistry required to ensure effectiveness (McGuinness, 2003; McKinstery & Topping, 2003; Topping 2002). This will form the focus of future studies with evaluation necessitated to determine the relative value of core components within cognitive instruction. The growing interest in the area makes it of particular pertinence for ongoing EP scientific scrutiny. As McGuinness (1999) argues " *The idea of thinking classrooms and schools as thinking communities, requires further articulation and interpretation, and the research base is in its infancy*" (p.30). This is an area ripe for development and it is argued that educational psychologists should focus their efforts on helping teachers to translate theory into practice into the classroom.

Acknowledgements

The author would like to extend thanks to the pupils and parents alongside school and staff who participated in this project. The support of Midshire County Council is gratefully acknowledged especially the Head of Psychology and Assessment who made this research possible. Similarly the support of fellow researcher Janet Wallace is greatly appreciated as is that of colleagues Mike Connor, Jeff Matthews, Nick English and Jennie Burrow.

Furthermore the author would express sincere thanks to the advice and assistance of Professor Carol McGuinness. Carol provided an invaluable contribution and catalyst to the project on a number of levels in particular allowing the use of ACTS to parallel the sister project in Northern Ireland. Likewise the author expresses gratitude to the insightful expertise of professionals in the field including Professor. Philip Adey, (Kings College, London); Dr. Steve Higgins, (Newcastle Upon Tyne University); Anne Robertson, (London Borough of Hammersmith and Fulham); Dr. Keith Topping, (University of Dundee); Dr. Steve Trickey, (Clackmannanshire and University of Dundee) and Dr. Chris Watkins, (Institute of Education).

Finally the author would convey thanks to the team at UCL. This included Dr. Phil Reed; Dr. Simon Watt and Dr. Robin Murphy. Warmest wishes are also given to Dr. Sean Cameron who provided the support and constant inspiration to illuminate even the darkest moments of real life research. The challenge and critical friendship was greatly appreciated. Thanks are also extended to the input from Dr. Steve Huggett throughout the study.

Chapter 1: Literature Review

Introduction

The purpose of this review is to identify and evaluate current thinking skills approaches, which aim to develop children's thinking to a qualitatively higher level. Past and present research will be considered with theoretical frameworks, research design, impact and limitations highlighted. Such a critique will then inform the foundations of future research in the realm of thinking skills.

Definitions

As a range of individuals demonstrate the term thinking skills is ambiguous and includes a range of processes (Higgins et al., 2004; McGuinness, 1999; Wilson, 2000). As Coles (1993) details thinking is "*a vast and intricate family of activities*" (p.335). Furthermore Wilson (2000) demonstrates the way in which there is no real consensus as to the inherent skills involved. For the purposes of this research the term will include the specific mental processes involved in thinking alongside the skills identified in more general taxonomies of thinking. The following definition encapsulates these concepts:

> *'The set of basic and advanced skills and sub skills that govern a person's mental processes. These skills consist of knowledge, dispositions and metacognitive operations.'* (Alvino, 1990) (p.50).

The rationale for such a descriptor is manifold to include a number of different concepts. Firstly it considers thinking as a set of skills and

processes rather than a knowledge base alone (Coles, 1993). Likewise it encompasses the range of '*higher order*' thinking skills such as critical thinking, decision-making and problem-solving highlighted in general taxonomies of thinking (Cotton, 1991; McGuinness, 1999). Moreover the term incorporates the concept of metacognitive aspects i.e. the capacity to reflect upon thinking and the processes involved (Topping, 2002). Finally it highlights the role of dispositions, which have been evidenced as critical in the development of thinking and reasoning skills in children (McGuinness, 1999).

Individual theoretical perspectives linked to specific thinking skills programmes will be scrutinised in the literature review. Such thinking skills approaches have arguably been of three types, principally those that focus on general thinking processes (bolt- on), those which are subject or domain specific (embedded) and those infused across curriculum areas (infusion) (McGuinness, 1999; McKinstery & Topping, 2003; Topping, 2002; Wilson, 2000). These programmes are identified in the DfEE research report number 115 as the most established and structured methodologies for developing thinking skills in the UK (DfEE, 1999). In addition a range of research reviews indicates they are the most researched and evaluated approaches (Connor, 2002; Higgins et al., 2004; McGuinness, 1999; McKinstery & Topping, 2003; Topping, 2002; Wilson, 2000).

Rationale for Research

The rationale for the research relates to the current interest in thinking skills in the domains of psychology, education, industry and politics (DeCorte, 2002; McKinstery & Topping, 2003; Topping, 2002; TES, 2002). At a societal level thinking skills are considered critical given the demands of modern life and technological advancements which demand '*higher order*' thinking (Cotton, 1991; De Corte, 2002; Gorodetsky et al., 2002; Wilson, 2000). As Gough (1991) argues '*specific knowledge will not be as important to tomorrow's workers and citizens as the ability to learn and make sense of new information'* (in Cotton, 1991) (p.1). Such societal changes have prompted governmental action in the area from research reports commissioned by the Department of Education and Employment (DfEE, 1999) to incorporating thinking skills as an explicit part of the National Curriculum. As Watkins et al (2001) point out; it now appears that '*learning about learning stands in its own right as a key goal for schooling*' (p.7). Indeed the teaching of thinking is now recognised as one of the primary aims of education (McGuiness & Nisbet, 1991; Mercer & Wegerif, 1999; Pithers & Soders, 2000). This is evidenced by the recent explosion in thinking skills packages available to schools (Adey, 2002; TES, 2002). Such approaches, however, are often not based on thorough evaluation (Gorodetsky et al., 2002). Indeed as Wilson (2000) describes '*evaluation studies are inconclusive*' (p.39). Such issues are a concern given the need for empirically supported interventions in education (Stoiber & Kratochwill, 2000). A very real need exists for an empirical evidence base in this field on both a national and international level (Bond, 2002; DeCorte, 2002).

This research aims to evaluate a teacher directed thinking skills approach in everyday classroom contexts with a clear research methodology. This has been highlighted as a need in the domain of psychology, as a research area calling out for more robust and generalisable evidence (Gorodetsky et al., 2002; Higgins, 2002; McGuinness, 1999; McKinstery & Topping, 2003; Watkins, 2003). A range of psychologists highlight the need to pinpoint key effective thinking skills packages (McGuinness, 1999; Watkins et al., 2001). The research aims to add to the current knowledge base for as the DfEE report highlights '*considerable evaluation work remains to be done in order to link the critical features of the framework to learning outcomes* '(p.28). This work aims to provide a scientifically rigorous evaluation to address the untouched issue of learner outcomes, resolve limitations of previous research and inform empirical evidence in this sphere. In such a manner a unique contribution will be made to the psychological research base in the longitudinal evaluation of a thinking skills intervention.

Overview

This review is divided into five key sections, the first of which identifies the theoretical perspectives and current key approaches for developing thinking skills. A critique of such programmes then follows with the implications and issues of past and present research documented. Future research options in general are then focused on followed by the forward directions. In particular the distinct contribution to psychology is described and hypothesis delineated in relation to previous findings. Finally a summary of thinking skills research past, present and future will be provided with the principal points reiterated.

Literature Review of Thinking Skills Research: Past and Present

Approaches to Developing Thinking Skills

The key thinking skill approaches can be distinguished between those which teach thinking as a separate skill in addition to the curriculum and those which develop thinking in specific or general areas of the existing syllabus (see Figure 1 for an overview) (Coles, 1993; McGuinness, 1999; McKinstery & Topping, 2003; Wilson, 2000). Interventions, which are aimed towards generic thinking skills, aim to teach specific component processes through specially designed programmes in a separate lesson, for example Feuerstein's Instrumental Enrichment (FIE) or Blagg's (1991) Somerset Thinking Skills. This general bolt-on approach argues that cognitive development enhanced in such a manner will impact across a range of contexts (Blagg, 1991; Feuerstein et al., 1980).

In contrast other structured approaches teach thinking in context, in a subject-specific area. This ensures thinking development is connected and embedded to the specific knowledge base and subject domain; examples include Cognitive Acceleration through Science Education (CASE) or Cognitive Acceleration through Mathematics (CAME) (Adey & Shayer, 1994) or Thinking through Geography (Leat, 1999). Finally thinking can be taught across the established curriculum using an infusion approach for example Activating Children's Thinking Skills (ACTS) (McGuinness et al., 1997). This aims to embed the teaching of thinking in all subjects so thinking skills permeate all aspects of a student's academic life.

These three approaches are those clearly delineated in research reviews (Connor, 2002; McGuinness, 1999; McKinstery & Topping, 2003; Topping, 2002; Wilson, 2000). Moreover they are the most notably researched and evaluated in the UK and most aptly represent the current national context (Connor, 2002; DfEE, 1999; Wilson, 2000). These three approaches were, therefore, selected for scrutiny in the literature review, which follows.

Theoretical Conceptions of Thinking

Distinctions can be made between the three key thinking skills approaches according to the theoretical foundations on which they are based. These are detailed in Figure 1. As McGuinness's (1993) review highlights the different conceptions of thinking and learning have led to alternative routes to cognitive instruction in teaching thinking. Structured generic thinking skill approaches aim to develop general thinking skills. They are based on the notion that thinking skills taught separately and in isolation can be transferred across contexts and applied over a range of situations. Such an orientation is founded on the various information-processing theories in cognitive psychology (McGuinness, 1993). These approaches identify a range of higher order thinking skills and design instructional materials to develop these specific processes. Feuerstein et al's (1980) Instrumental Enrichment is a classic example of this. This programme concentrates on a set of cognitive processes, which are then developed through specific materials. The concept of 'mediated learning' is crucial, in which learners are shown a particular cognitive method for approaching a problem in a given scenario.

The key issue for such approaches is the extent to which cognitive processes are transferred across settings (Blagg, 1991; McKinstery & Topping, 2003).

The subject-specific approach to teaching thinking draws directly from Piagetian frameworks with the idea that knowledge is based on interactions with the environment and mental representations (Piaget, 1972). In particular the notion of cognitive conflict is central in aiding development and prompting change in mental constructs. Similarly Vygotsky's (1978) theory of the zone of proximal development with the role of mediation supporting and forwarding cognitive development is inherent in such approaches. One such example would be CASE (Cognitive Acceleration through Science Education), which seeks to increase student's thinking to formal operational level (Adey & Shayer, 1993; 1994) as delineated by Inhelder and Piaget (1958). The intervention aims to create cognitive conflict and higher order thinking in a subject-specific domain. Such a methodology ensures thinking skills and conceptual knowledge develop simultaneously in a specific area to accelerate cognitive development (Adey et al., 2002). Critical problems, however, rise out of the debate regarding transference to other curriculum areas and application generally (Coles, 1993; McGuinness, 1999; Perkins & Grotzer, 1997). As Coles (1993) argues the approach "*does not enable children to forge links between different areas of knowledge*" (p.339).

Finally the infusion approach to thinking skills intends to embed thinking skills within the existing curriculum. It allows for curricular content and thinking skills to be inextricably linked being both taught and learnt together. Its basic

cognitive framework is founded on taxonomies of thinking skills (Swartz & Parks, 1994), which are then mapped onto curriculum contexts. Infusion approaches are clearly evident in practice from Accelerated Learning (Smith, 1996) to Activating Children's Thinking Skills (ACTS) (McGuinness, 1999, 2000). Despite this, however, the impact and effectiveness of such programmes remains to be evidenced in more detail (DfEE, 1999; McGuinness, 1999; Wilson, 2000).

Figure 1 : An Overview of the Different Thinking Skills Approaches and Empirical Evidence Base in the United Kingdom

Generic Thinking Skills taught in context free situation. **Bolt- on Approach**	IE (Instrumental Enrichment) (Feuerstein, 1980; Blagg, 1991; Savell et al., 1986).	Thinking as Information Processing (Simon, 1979; Kail & Bisanz, 1992).	Two by two pre-post design over 2 years. Issues of inappropriate measures and variables with no reasonable control.	Basic analysis and conclusions indicated no improvement in FIE trained groups in UK.	Clear conceptualisations and experimental design. Although some limitations of measures and variable categories.
Subject specific intervention to target thinking in certain curriculum areas. **Embedded Approach**	CASE (Cognitive Acceleration Through Science Education) (Adey & Shayer, 1993, 1994; Adey, 1997; Shayer, 1996).	Thinking as Sense Making (Piaget, 1972; Vygotsky, 1972).	Two by two pre-post design over 2 years, extended later to 5 years. Concerns over lack of control group and extraneous variables unaccounted for.	Mixed results with effects unstable over time and gender. Conclusions not necessarily consistent with data.	Strong theoretical underpinnings with large scale and scope. Issues with research design.
Infusion methodology teaching thinking across the curriculum. **Infusion Approach**	ACTS (Activating Children's Thinking Skills) (McGuinness et al., 1997, 1999; Swartz & Parks, 1994).	Thinking Across Contexts (Swartz & Parks, 1994).	Exploratory study to investigate student and staff reactions to ACTS. Posttest of teachers only. No clear research design with adequate controls.	Perceived positive impact, however, measures lacked reliability and validity. Limitations were identified. Improvements highlighted.	Useful basis for follow-up research. Key difficulties relate to research design and confounds.

Literature Critique: Issues and Implications of Research Past and Present

The existing research base appears to suggest that thinking skills instruction has a positive impact at a student, staff and school level (DfEE, 1999; Higgins et al., 2004; Watkins et al., 2001). This section aims to examine such promising effects and determine the extent to which they are evident in the psychological literature. Firstly specific issues and findings related to the empirical evidence will be addressed. More general matters will follow, as identified by research in the area with implications for future development highlighted.

Specific Issues and Implications of Research Past and Present

Student's Thinking Skills

A range of current thinking skill programmes propose to enhance and develop children's thinking skills to a qualitatively higher level (Adey et al., 2002; Adey & Shayer, 1993; Lipman, 1991; McGuinness, 1999). Indeed some studies (Feuerstein et al., 1980) do evidence differences in students reasoning skills. It is not until closer analysis, however, that this data is called into question. Moreover the actual range of studies do not unequivocally demonstrate enhanced student thinking skills. Instead both past and current research reveals a mixed picture of whether changes in higher order thinking processes actually occur (DfEE, 1999).

The programme heralded as the most established and evaluated thinking skills approach is Cognitive Acceleration through Science Education (CASE) (McGuinness, 1999; Topping, 2002; Wilson, 2000). The main objective of CASE was to move children's' thinking to the level of formal operations thereby developing cognitive processes. The research claimed to enhance children's thinking and this was measured by a Piagetian reasoning task. In terms of students' thinking immediate results following intervention indicated a change in reasoning of the 12 year old boys' cohort, but not others. Moreover data reported on later (Shayer & Adey, 1993) illustrated gains in reasoning, however, this was due to a minority of students showing very large gains rather than general effect. Likewise at the final post test (Shayer & Adey, 1993) the data reported focused on science achievement tests and GCSEs alone. No information was provided on an impact on the students thinking skills. This does not provide conclusive evidence as to the effectiveness of CASE in developing thinking skills specifically. Additionally the measure used to ascertain thinking skills was not described in relation to reliability or validity statistics. It was designed to assess '*reasoning*', which acts as only one aspect of the myriad of thinking skills and would need a tighter definition in future. No pilot study to gauge its effectiveness was conducted which limits its potential ability to tap into thinking. Finally reviews of such methods also raised concerns about the transfer of skills to other domains (Perkins & Grotzer, 1997).

More recent research into cognitive acceleration, with younger children, has evidenced significant immediate effects on children's rate of cognitive

development (Adey et al., 2002). Long term evaluation is required and is acknowledged by the researchers as crucial in ascertaining the longevity of effects (Adey et al., 2002). The measurement devices utilised in the study were also very specific cognitive development tasks based on original Piagetian protocols (Inhelder & Piaget, 1958). Such techniques have been critiqued in terms of the ability to accurately ascertain children's skills (Dasen, 1972; Donaldson, 1978). Research generally in the area of cognitive acceleration has been limited by the use of such specific devices.

Another approach, which claims to enhance student's thinking skills, is ACTS, Activating Children's Thinking Skills (McGuinness et al., 1997). Research reported benefits for children's thinking including; greater reasoning and creativity, an improved ability of children to clarify their thinking processes and for children to become more focused in their approach to thinking. Such positive effects, however, were identified by the teachers involved in the project, rather than any measure of thinking with the students. The impact on students thinking was, therefore, assessed not by an objective or standardised measure of performance but rather by staff's subjective perceptions. This introduces a substantial degree of bias into the proceedings. As McGuinness (1999) herself pointed out '*the data are teachers impressions and opinions*' and cannot arguably represent definite changes in students thinking conclusively. Moreover the measure used to gauge these opinions had not been standardised prior to the project with no validity or reliability data reported. Likewise no pre test occurred and the

open-ended questionnaire was used at the end of the study, which may also have affected results.

Research investigating the peer tutoring of thinking skills has, likewise, evidenced positive effects in the domain of thinking skills for both tutors and tutees (McKinstery & Topping, 2003; Topping & Bryce, 2004). At present, however, like the McGuinness (1999) study this is indicated by subjective participant feedback alone as opposed to measures of cognitive development. As McGuinness (1999) recognises substantial evaluation remains.

Other studies have also had difficulties in outlining a change in students thinking skills. The evaluations of Instrumental Enrichment demonstrate a mixed view. Savell et al (1986) reported improvements in standards on non-verbal measures across a range of studies. However these evaluations had only ever been conducted at the direct end of the intervention with no follow up measures (DfEE, 1999; Shayer & Beasley, 1987). The UK evaluation study by Blagg (1991) indicated no significant improvements between experimental and control groups. Both studies were, however, limited by the lack of a reliable and valid measure to tap into students thinking skills. As Blagg (1991) argues '*there is a desperate need for the development of new, reliable and valid procedures that can sensitively assess and monitor attitudinal, behavioural and cognitive changes in both pupils and adults*' (p.141).

The key issue drawn from the research is the lack of any hard, unequivocal evidence to prove such packages enhance students thinking skills. As Wilson (2000) comments, "*given the paucity of evidence, it would, perhaps, be fairer to conclude, "the jury is still out" on this particular issue*" (p.39). This is linked in part to the lack of any standardised measure to ascertain thinking skills. Future research will need to devise and use measures, which tap into children's thinking skills more readily. In particular there is a need to focus on the learner in terms of thinking as well as the attitudes, behaviours and perceptions linked to such processes. The area of social and affective aspects of thinking is increasingly heralded as an area worthy of further exploration (DfEE, 1999; McGuinness & Nisbet, 1991; Watkins et al., 2001; Wilson, 2000).

Student Attainment

The main impetus behind thinking skills approaches is the premise that they will raise standards and attainments of students (Adey et al., 2002; DfEE, 1999). Although evidence exists to illustrate improved attainment there are restrictions as to the conclusions that can be drawn. For example, the CASE projects demonstrated enhanced GCSE results of one grade in Science and smaller gains in Maths and English for students. Adey and Shayer (1993) identify how this demonstrates that even after the intervention the effects of CASE continued. Such an assumption is not seen in the data, however, as specific groups of students did not show gains, with effects being unstable over time and gender (Watkins et al., 2001.) Likewise the research design limits the conclusion that CASE was the prime mechanism for such an

impact. The lack of a reasonable control meant children in the *experimental* group were compared with a *no treatment control*. The impact of the project may not, therefore, necessarily be CASE itself but rather another variable such as talk time in class or the expectancy of change. An *attention placebo* or *waiting list control* could have eradicated such issues and supported the positive impact of CASE more conclusively. Recent studies of cognitive acceleration with younger children did use a control group and used the teachers to baseline test student's attainments using the "signposts" framework (Birmingham, 1997). As Adey et al (2002) point out, however, assessments "*were administered by teachers with little special training in their administration, and without independent verification*" (p.7). This highlights the limitations of research in gauging objective information on student attainment.

Difficulties with control groups also limit conclusions drawn in other studies. In the Philosophy for Children study (Lipman, 1991), although improvements in student's reading and maths occurred, no detail is given on the control groups. This omission limits the conclusions that can be made as a range confounds exist in the design. Similarly the ACTS project (McGuinness et al., 1997) reported a perceived positive impact on student attainment. The research design, however, used no measures of student outcomes or any control group as means of comparison. Instead the data illustrated staff perceptions of student change. Such data does not accurately assess or gauge students' attainment but merely teachers' opinions on it.

Within the studies of Instrumental Enrichment the evidence regarding student attainment is contradictory. Savell et al (1986) indicated substantial significant effects on attainment as measured by non-verbal measures of intelligence. However, the impact on other attainment and achievement measures were inconsistent. Blagg's study (1991) in the UK illustrated no significant improvements on intellectual performance or tests of reading or maths skills on students who received the intervention.

Research in the domain of peer tutoring of thinking skills has yielded positive student outcomes, namely in reading skill development (McKinstery & Topping, 2003; Topping & Bryce, 2004). In particular studies illustrate reciprocal tutoring of students produce gains in reading comprehension (Fuchs et al., 1999; Simmons et al., 1994). This effect was noticed when a paired thinking method was adopted with an additional impact on thinking skills (McKinstery & Topping, 2003; Topping & Bryce, 2004). The latter effect, however, was evident in subjective participant feedback alone and indicates the need to obtain more rigorous data on the impact of interventions.

In summary, the evidence regarding improved attainment is contradictory. Likewise the research design and measures used often limits the validity of the conclusions drawn. There is a need for more robust and generalisable results as to the effect on student attainments. The need for data to inform practice has been highlighted as critical to determine the value of thinking skill packages on students (Gorodetsky et al., 2002; Higgins, 2002; Watkins et al., 2001). As Stoiber and Kratochwill (2000) point out '*practitioners can*

no longer assume that offering something is useful; there is an expectation of professional accountability in determining students needs and linking them to optional preventative and intervention strategies' (p.100). Future research needs to provide evidence that thinking skills packages do have a positive impact on student outcomes both in terms of thinking and attainment.

Staff Effectiveness

There is growing research, which indicates the application of thinking skill approaches have a positive impact on enhancing teacher and thereby school effectiveness (McGuinness, 1999; McKinstery & Topping, 2003; Munro, 1999; Stewart & Smardon, 2002; Watkins et al., 2001). A study in Australia with 32 secondary teachers indicated that teachers who engaged in a professional development programme to learn about learning resulted in enhanced teacher effectiveness (Munro, 1999). The findings illustrated an increase in the display of effective teaching behaviours, which were maintained over the two terms following the programme. Moreover teachers demonstrated an increased use of strategies consistent with learning theories. Likewise teachers reported that their knowledge and beliefs about learning had improved, thereby enhancing their delivery of classroom teaching. Despite such positive findings the study's main limitation was the selection of teachers for the project. Although the researchers sampled schools, the individual teachers involved were self-selected. This could act as a confound for the research as it may be only teachers interested and receptive to change engaged in the programme. This limits the validity of the conclusions drawn to this self-selecting sample.

Such criticisms can also be levelled at Blagg's (1991) evaluation of Instrumental Enrichment. Although some significant results did emerge in terms of developing teacher's positive attitudes to students the teacher sample was self selected which may have impacted on results. Moreover the teacher measures included a range of scales, which "*were not factor analysed or checked for validity or reliability*" (Blagg, 1991) (p.45).

Evidence from the ACTS project in Northern Ireland, from the teachers directly involved, demonstrated a positive impact. Staff felt the project clarified their concept of thinking skills and enabled more effective planning (McGuinness, 1999; 2000). The use of ACTS was also reported to have a positive impact on staff's conceptual knowledge and skills in this domain. These findings, however, are staff's perceptions as measured by an open-ended questionnaire. They are not substantiated or triangulated with more quantitative information relating to knowledge or application of skills. They offer, therefore, the staff's perceptions only of positive change. The issue of accurate measurement of change is evident in other studies. A study in New Zealand illustrated improved teaching when a focus on critical thinking had been adopted by staff (Stewart & Smardon, 2002). The results, however, were based on the teacher's and researcher's descriptions as opposed to any objective measure. Similarly work on Paired Thinking demonstrates developments in thinking and reading skills for tutors alongside tutees (McKinstery & Topping, 2003; Topping & Bryce, 2004). At present, however, this too is evidenced by only qualitative feedback.

Although, evidence initially suggests enhanced teacher effectiveness, this is not necessarily the case after closer analysis. The limitations of both self-selecting samples and lack of objective measures have restricted conclusions drawn in this domain. Indeed some studies did not even address teacher change such as CASE (Adey & Shayer, 1993) or Let's Think (Adey et al., 2002). There is a need for more objectively selected samples and more external quantitative measures of change in the area of staff development.

General Issues and Implications of Research Past and Present

Research Design Issues

Most of the existing research into thinking skills is based on small self-selected samples of teachers and pupils often under ideal conditions. Projects including ACTS, IE, CASE and Let's Think all involved the element of staff self-selection, which may have limited the conclusions drawn from the research. Despite the fact schools were randomly selected, the staff that actually engaged in the projects volunteered to participate. The main issue with self-selection is it is unknown along which psychological dimension it operates and is, therefore, difficult to control for. As the TES (2002) highlights '*the problem is that almost all thinking research has measured the success of highly motivated teachers trained and monitored by experts*' (p.25). Indeed '*the impact on ordinary classrooms is very small*' (McGuinness, 1999) (p.9). Research into everyday classroom contexts with the broad spectrum of the teaching population catered for is now required. As Wilson (2000) points out '*most research has been conducted under*

optimal learning conditions' (p.39). A further implication for future research is the evaluation of thinking skills approaches in everyday practice in a range of schools with a cross section of randomly selected staff. As McGuinness (1999) highlights "*problems with scaling up and transferring the effects to everyday classrooms have been identified*" (p.29.)

Current research has also been limited by inappropriate research design which has meant a range of potentially confounding variables have been unaccounted for. Demographic characteristics of student, staff and school populations are generally not provided in the literature (Higgins et al., 2004). These factors are not addressed let alone controlled for in the research design, with the impact of gender, culture, geographical location and social economic status largely ignored. Future work will need to detail such factors so there possible role as confounding variables can be minimised.

The lack of reasonable control groups in previous research is another limitation of experimental design. Indeed this is a longstanding issue as demonstrated by Sternberg and Bhana (1986) in a review of thinking skill interventions; " *most studies lacked control groups or used inadequate controls*" (cited in Blagg, 1991) (p. 140.) Prior studies have all used in effect a *no treatment* control group to compare against the *experimental* group receiving the thinking skills input. In the instance of CASE (Adey & Shayer, 1993) such a design was used which meant extraneous variables were not adequately controlled for. Future studies will need more rigorous use of a reasonable control so as to discredit potentially confounding variables.

Moreover the way in which samples in the *control* and *experimental* group were matched in previous studies was often inappropriate. In Blagg's (1991) evaluation of Instrumental Enrichment for example the matching occurred on an ad hoc basis linked to teacher's decisions. More careful matching of *control* and *experimental* groups will be required in future research to allow for meaningful comparisons to be made both between and within groups. In addition more description of the control groups will be necessary for as Lipman et al's (1980) study on Philosophy For Children indicated results were "*difficult to assess because of the lack of detail provided on the control groups*" (DfEE, 1999) (p. 14). New work in this area will warrant the use of more effective control groups with demographic characteristics specified.

Finally in relation to research design, previous projects have been restricted by the lack of detail in relation to measures. As Gorodetsky et al (2002) highlight, few studies use thorough forms of evaluation. Connor (2002) cites Bradley (1983) and Burden (1987) who criticised studies in identifying reliable measures of outcome, which were statistically sound. Indeed the major projects looking into ACTS, CASE, IE, Let's Think and Philosophy for Children provide no data on the reliability and validity of the measures used. Likewise the lack of any pilot studies in previous work reduces the rigour of design. The evaluation of ACTS used no pre-measure but an 18 item open-ended questionnaire constructed especially for the project. The population on which the measure was sampled and standardised is not detailed. Furthermore in Blagg's (1991) evaluation of Instrumental Enrichment some measures were indirect subjective teacher ratings of student's performance

as opposed to standardised measures. As Blagg (1991) points out "*researchers are posed with the dilemma of using inappropriate, standardized procedures with known characteristics or newly designed procedures, inadequately researched and therefore of unknown reliability and validity*" (p.38).

Previous research has provided little justification as to measures selected and their rigour. Unfortunately it appears little has changed since Sternberg and Bhana's 1986 review of such programmes which detailed "*the kind of outcome measures used to support the programmes were wholly inadequate*" (in Blagg, 1991) (p.140). Future research will need to provide explicit reasoning for measures selected alongside the technical detail and piloting to evidence sufficient rationale for use. As Watkins et al (2001) point out '*the choice of performance measures and whether they assess high level learning will be critical*' (p.2).

Unresolved Issues

A key issue, which is inherent in any teaching package, is transfer - the extent to which skills taught are later applied beyond the context of acquisition. The critical importance of this issue has consistently been at the centre of the thinking skills debate (DfEE, 1999; Nisbet, 1993; McKinstery & Topping, 2003; Wilson, 2000). As Watkins et al (2001) highlight this is the '*crucial element*' and yet it is often un-addressed in the prevailing research literature. Incorporating the teaching of transfer in some approaches is specified, such as ACTS in the form of near and far transfer in the plenary of

thinking lessons. In general, however, the issue is not clearly tackled. As Sternberg and Bhana's (1986) review of cognitive intervention studies indicated; "*inadequate attention was usually given both to the transfer of training and to durability of training over the long term* "(in Blagg, 1991) (p.x). It seems these issues still persist with transfer omitted or ignored in research. Indeed CASE, Let's Think, Instrumental Enrichment all offer isolated programmes, which do not build the transferability of skills to real world contexts. Further work will need to recognise and accommodate the importance of transfer with more empirical evidence.

Practical issues linked to embedding thinking skill approaches into the curriculum are also unresolved in the literature (McGuinness, 1999; TES, 2002; Topping, 2002). As McGuinness (1999) highlights the practical implications to classrooms in terms of curricular design, materials and pedagogy are potentially huge. Future research would need to explore such effects more closely to ascertain the potential wider impact.

Finally a review of the research literature indicates the range of methodological limitations present in previous work. As Wilson (2000) highlights '*evaluation studies are inconclusive*' (p.39) with a real need for more empirical evidence in this area. At present the key issue is that research does not substantiate fully whether thinking skills approaches do have a positive impact. As Coles (1993) argues '*the blunt statement "it is possible to teach thinking" is an empirical one and so ought to be backed by empirical evidence*' (p.341). Future research needs clear rigorous

methodologies with application in a range of settings to produce robust and generalisable evidence (Gorodetsky et al., 2002; Higgins et al., 2004). Indeed further detailed analysis and controlled evaluations are required (Deutsch, 1997 in Connor, 2002; McGuinness, 1999; Watkins et al., 2001). Moreover the use of theory driven research is needed with links between practice and research clear (Bond, 2002; De Corte, 2002). Clearly as McGuinness (1999) argues, there is a need for '*considerably more systematic work in this area*' (p.29).

Summary

Current psychological literature, although indicative of some positive effects in relation to thinking skills, is limited by a range of methodological and measurement errors. As Coles (1993) highlights "*the variety of teaching thinking programmes might be criticised as a confusion of educational aims and curriculum agendas*" (p.333). Moreover Blagg (1991) details how "*close scrutiny of the research reveals many inadequacies, methodological flaws, and over-optimistic interpretations"* (p.25). General research design issues of unsubstantiated measures and lack of reasonable controls, means that conclusions are weakened. The assumed specific positive impact on staff and students alike is not borne out by statistical analysis or proven by the measures utilised. Little research attempts to measure the changes in children's' thinking and the transfer of these skills into alternative contexts. Likewise the lack of demographic characteristics in the research limits its generalisation to the population as a whole (Higgins et al., 2004). In future work, studies of a more scientifically rigorous nature are needed to generate

empirical evidence to explore the effectiveness of thinking skills programmes (DeCorte, 2002; Gorodetsky et al., 2002; Higgins et al., 2004; McKinstery & Topping, 2003; Wilson, 2000).

Forward Directions for Future Research

From a review of the literature it appears that, although, a number of studies exists many of these are methodologically unsound. This piece of research aims to resolve the earlier limitations of research past and present. The following section outlines how such difficulties will be addressed whilst also providing a rationale for this particular study. Moreover the unique contribution to the psychological research base will be highlighted with research questions and hypothesis illustrated.

Resolving Limitations of Research Past and Present

Earlier work in the domain of thinking skills has been limited by inappropriate methodological design and a limited size and scope, which have reduced its impact on developing the empirical evidence base. As McKinstery and Topping (2003) point out, although teacher directed thinking skills interventions are widely used they are *"not well or positively evaluated"* (p.202). This research addresses such criticisms in designing a scientifically rigorous study to inform the foundations of theory, research and practice. The following key issues have been highlighted as areas of difficulty in previous research, which this study seeks to remedy.

Methodological Issues

With regards to the measures used in previous research, few studies used tools to aptly evidence the effects of thinking skills approaches on children. Indeed there was a restricted focus on attainment measures only with the general area of other learner outcomes ignored. As McGuinness (1999) noted in the DfEE review of research '*considerable evaluation remains to be done in order to link the critical factors of the framework to learning outcomes*' (p.1). In particular there was a distinct lack of data regarding the thinking skill development of students. The only studies, which actually used a measurement device, was the CASE and Let's Think approaches, which used Piagetian reasoning tasks. No other piece of research has attempted to directly tap into children's thinking skills explicitly. This research aims to explore developments in children's' thinking over time and in relation to the intervention.

Previous research tended, as Wilson (2000) points out to focus exclusively on attainment alone. Similarly such measures were often used at the end of the project period such as the GCSEs in the CASE approach. This project aims to use measures of attainment, which could occur over time to establish development and possible transfer of thinking skills development. This could use a system common to all schools and a pertinent external measure.

In the current research base there is a lack of information on children's perceptions about their own thinking and learning. This is in direct contrast to the '*growing recognition of the importance of affective factors in thinking-*

attitudes, motivation and disposition' (McGuinness & Nisbet, 1991) (p.181). Indeed the juxtaposition between present evidence and the professional weight placed on affective factors and metacognition is huge. A range of psychologists have highlighted the role of perceptions and emotions as crucial in the development of thinking (De Corte, 2002; McGuinness, 1999; Watkins et al., 2001; Wilson, 2000). Despite this sentiment, however, no research has examined student beliefs about thinking or learning. This piece of research aims to rectify this by exploring student self-perceptions.

Also in relation to measurement issues this study seeks to provide a rationale of measurement device selection. Earlier research provided little explicit reasoning regarding the choice of measures used. As Sternberg and Bhana (1986) also describe, "*outcome measures often overlapped program content, and thus tended to favour the program being tested*" (in Blagg, 1991) (p.x). Moreover little detail was given in terms of validity, reliability or standardisation figures. This study aims to redress both these issues.

Previous studies have tended to use a *no treatment control* as a comparison with the thinking skills intervention group. Such a control meant a range of factors, other than the programme itself could have generated the effects. This research will use a more effective control group to ensure a more rigorous research design. Moreover samples within the *control* and *experimental* group will be appropriately matched rather than on the ad hoc basis present in earlier research.

Finally earlier research tended to have a small sample restricted to one geographical location with no demographic characteristics specified. This meant it was difficult to gauge how representative the sample was of the national and international population. This research piece proposes to use a range of locations and schools, sampled across a number of criteria from social economic status to SATs. Such a design should make the study representative of the current situation in the UK.

Size and Scope Issues

A number of criticisms have been levelled at earlier research with regards to its limited size and scope (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000). As Wilson (2000) highlights ' *problems of embedding the approach into everyday classroom practice, based upon what the average teacher can achieve rather than the expert practitioner working in good conditions with well motivated pupils, remains to be demonstrated* " (p.39). Indeed in all previous work teachers have volunteered to participate in the programmes. Such self-selection may have meant only highly motivated and committed teachers to developing thinking skills were involved. This research aims to include a range of everyday classroom contexts with teachers not self-selected but selected by alternative criteria. To some extent this could go some of the way in answering the question posed by the TES (2002) '*How can we ensure good teaching of thinking by the ordinary teacher in the average school?*' (p.25).

Restricted Empirical Evidence Base

At present the research base provides an ambiguous and often mixed picture in terms of the impact of thinking skill packages. Given the potentially vast costs both financially and time wise of such approaches there is a need to establish if they are both valuable and effective for students, staff and schools. As Gorodetsky et al (2002) argue there is a need for more systematic evaluation of all the new techniques appearing. In particular as Bond (2002) outlines, studies will need to demonstrate the benefits. If this is not done then this sphere of knowledge will be dismissed as evidence is lacking. Research of a more scientific nature is required as highlighted by both the DfEE report (1999) and psychologists in this area alike (De Corte, 2002; Higgins et al., 2004; McGuinness, 2002; McKinstery & Topping, 2003; Topping & Bryce, 2004; Wilson, 2000). As De Corte (2002) suggests the flaw of earlier research lies in its small experimental groups and poor design. This research aims to use larger experimental groups matched with appropriate controls to derive more reliable and generalisable conclusions.

In particular to the thinking skills research base the infusion methodology has, unlike other approaches had little rigorous evaluation. To date the only evaluation of such an approach was in Northern Ireland with the ACTS project with teacher only measures (McGuinness et al., 1997). No scientific study has measured the impact of ACTS on learner outcomes in ordinary classrooms in the UK. The infusion methodology is the most under researched domain of thinking skills at present, making it an area of future exploration through scientifically rigorous research.

Research Aims and Questions

From the conclusions drawn from previous research issues there are several aims, which this research would aim to fulfil. These are as follows and linked to resolving earlier limitations

Methodological Issues

1. To examine the effect of thinking skills approaches on a range of learner outcomes namely; thinking, attainment, internal self perceptions and external behaviours.

2. To justify and explain the use of specific measurement tools to tap into learner outcomes with reference to reliability and validity.

Size and Scope Issues

3. To explore the impact and possible use of a thinking skills approach in ordinary classrooms with a cross section of staff and schools.

Restricted Empirical Evidence Base Issues

4. To use a scientifically rigorous design with appropriate control group and a representative sample of the population.

5. To evaluate the impact of an infusion methodology (ACTS) in children over varying durations of time to determine if one year intervention is as effective as a two year intervention period.

These overarching aims would generate the following research questions:

1. What are the effects of a thinking skills approach, namely ACTS, on children as learners, especially their thinking, attainment, self-perception and behaviour?

2. How can we establish a link between thinking skills and learner outcomes in particular related to thinking, attainment, self-perception and behaviour?

3. Can a thinking skills intervention, such as ACTS, be used in ordinary classrooms in a range of different schools with a variety of teachers?

4. Does a thinking skills intervention, such as ACTS, have any effect on students when compared to an appropriate control group in terms of thinking, attainment, self-perception and behaviour?

5. What impact does the ACTS (infusion approach) have on children across varying periods of intervention duration? Are the effects transferred across settings and maintained over time?

Rationale for Study

For the purposes of this study it was decided to concentrate on the impact of a thinking skills approach on primary aged learner outcomes. The reasons for this are numerous. Firstly the majority of research has used inappropriate measures to determine the impact on learners across contexts. Likewise it has failed to effectively assess the thinking skills of children effectively. This research should resolve such earlier methodological flaws by focusing explicitly on learner outcomes with a specific examination of thinking skills development. The research should also fill in the gaps of previous work by adopting a primary focus. Past research has tended to focus on secondary or higher education settings (McKinstery & Topping, 2003; Topping, 2002). This project will, therefore, apply the approach to a primary age range to inform the knowledge base in this area. In a similar fashion the context of ordinary classrooms from a range of schools will be utilised. It is hoped that such an approach will redress the limited size and scope of previous work whilst applying a thinking skills intervention to everyday classroom contexts.

The research proposed will additionally aim to inform the research base by using a 2-year intervention evaluation study of a thinking skills programme. Such a design builds on earlier work, which has used a two-by two-pre post method over such a period (Adey et al., 2002; Blagg, 1991; Feuerstein et al., 1980). This study intends to develop the empirical evidence base by using an effective control group with a representative and randomly selected sample. Moreover the work aims to develop the research base by exploring the issue of intervention duration by comparing one and two years of

intervention. This challenges the predominant view based on earlier work that a two-year period is considered appropriate in effecting long-term cognitive change (Adey & Shayer, 1993; Blagg, 1991; Feuerstein et al., 1980; McGuinness, 1999). This contrasts with more recent research findings suggesting a shorter time period is required in order to generate cognitive change in children (Adey et al., 2002; McKinstery & Topping, 2003).

For research purposes one approach in particular will be the focus as the thinking skills intervention - the infusion methodology, namely ACTS (Activating Children's Thinking Skills). The reasoning behind this is that there are no scientifically rigorous evaluations in this sphere of thinking skills. The majority of large-scale evaluation studies have occurred in other areas of thinking skills instruction from the subject-specific approach such as CASE to the generic approach of IE. Few studies have focused on the infusion methodology, with those that do exist acting as preliminary explorations into the approach. The research, therefore, aims to inform the empirical evidence base as to the effectiveness of an infusion methodology approach, namely ACTS. Initial work has commenced on this in Northern Ireland, however, this research was restricted to teacher perceptions only (McGuinness et al., 1997).

Another reason for selecting the infusion methodology as the thinking skill intervention was the fact it is the most easily transferable to the school context. Infusion seems to be the optimal operational technique to embed thinking skills into the everyday classroom curriculum. As such, infusion

does not require an additional lesson such as the generic thinking skill approach and therefore does not impede on curriculum time. Also by its very nature it is infused and permeates every aspect of the curriculum so the transfer of thinking skills is addressed. Infusion is, moreover, an ideal approach for the primary age range when children tend to have one class teacher for the year that can explicitly make links across curricular areas. McGuinness (1999) explains the additional benefits of infusion as; the optimal use of class time alongside teaching thinking directly matched to topics in the curriculum. This ensures content instruction is invigorated leading to deeper understanding. The project, therefore, aims to apply a rigorous research design in the sphere of the infusion methodology. The ACTS intervention in particular was chosen as initial research by McGuinness et al (1997, 1999) indicated positive outcomes. The research design, however, had no control group and no measures on learner outcomes. This piece aims to remedy earlier limitations by the use of a more scientific methodology.

Finally the rationale behind the study is in terms of the unique contribution to the psychological research base. The project aims to develop the existing evidence base in relation to the infusion methodology for teaching thinking skills. No scientifically rigorous evaluation of infusion has occurred as yet, only an initial exploration (McGuinness et al., 1997). This study will therefore, apply research to a new sphere of thinking skills and contribute to the empirical evidence base as to whether an infusion thinking skills package is effective. As both Wilson (2000) and Topping (2002) indicate, current

evaluation studies are inconclusive with research poorly designed and as McGuinness (1999) stresses '*considerable evaluation remains to be completed*'. Likewise, De Corte (2002) pinpoints the need for research with experimental and control groups to evidence impact clearly. This need to develop the theory and empirical evidence base is heralded as critical (Bond, 2002; De Corte, 2002; Higgins et al., 2002; McGuinness, 2002; McKinstery & Topping, 2003; Topping & Bryce, 2004; Topping; 2002, Watkins et al., 2001, Wilson, 2000). This piece of research would aim to achieve such ends by providing an intervention evaluation study of the impact of a thinking skills approach, thereby, contributing to the empirical foundations of this area.

The study will inform the psychological evidence base in relation to theories of cognitive development and intelligence. As McGuinness (1993) highlights '*the practice of teaching thinking heralds new signs and developments in cognitive theory*' (p.305). The research could ascertain if programmes could enhance students' thinking processes. They would add to the debate regarding multiple intelligence or fixed intelligence along with cognitive development theories (Connor, 2002; Watkins et al., 2001; Wilson, 2000). This would contribute to the development of cognitive theory in this sphere.

Hypothesis to be examined

The research aims to concentrate on the impact of thinking skills approaches on learners in particular. It was thought a focus on school and staff impact would be too wide for the purposes of this project. Previous studies indicated improved thinking skills of students, which prompt the question 'Does thinking

skills instruction actually alter children's thinking?' Evidence from subject specific and infusion methodologies suggest changes in children's thinking. The hypothesis therefore states:

1. *Students will have developed thinking skills having received the ACTS intervention compared to the control group.*

Earlier research has tended to focus on students attainment in relation to interventions. However, little work has looked into children's self-perceptions and dispositions to learning. This prompts the research question 'Does thinking skills instruction change a student's self-perception as a learner?' Such a question leads to the hypothesis:

2. *Students will demonstrate a change in self-perception of themselves as learners having received the ACTS intervention compared to the control group.*

A gap in earlier research relates to the effect of thinking skills interventions on behaviour on social situations. This generates the research question 'Does a thinking skill approach impact on a student's behaviour in social situations?' Such a linked hypothesis would be:

3. *Students will show a behavioural change in social situations having received the ACTS intervention compared to the control group.*

A range of research suggests possible improvements to students over varying periods of intervention. This demands the question 'Is a one year intervention of ACTS as effective as a two year intervention?' This would equate to the hypothesis:

4. *Students will make greater gains following two years of the ACTS intervention compared to one year of the ACTS intervention.*

Finally previous studies have indicated positive changes to staff in connection with thinking skills interventions. This generates the research question "Is there an effect on teacher practice and knowledge during the implementation of a thinking skills package?" This would connect to a hypothesis:

5. *Teachers will have a developed understanding and changed practice to the teaching of thinking skills following the implementation of the ACTS intervention.*

Conclusion

The literature review summarises the range of research into thinking skills approaches in the UK in the last 10 years. In particular, the research into subject-specific and generic thinking skills packages suggest a design to be used to examine the newer infusion methodology. Despite the variety of studies, a number of key issues and limitations restrict the conclusions drawn from earlier work. The methodological omissions of appropriate controls and

inappropriate measurement devices minimise previous findings. This project aims to resolve earlier difficulties and inform the research, by the use of a focused, scientifically rigorous intervention evaluation study. The intention is to develop the empirical foundations of thinking skills research whilst informing practice issues in everyday contexts. The next chapter details the way in which the research design proposes to overcome previous methodological flaws. Such an area has been highlighted as critical by a number of psychologists in this domain (Bond, 2002; De Corte, 2002; Higgins et al., 2004; McGuinness, 2002; McKinstery & Topping, 2003; Wilson, 2000).

Chapter 2: Research Methodology

The purpose of this study is to scientifically evaluate the impact of a thinking skills approach whilst, overcoming the limitations of earlier research. As Blagg (1991) argues *"close scrutiny of the research reveals many inadequacies, methodological flaws and over optimistic interpretations*" (p. 25). This research aims to address such criticisms and examine unanswered practice questions as informed by previous study and the preparatory phase (see Appendix 1 for detail). The preparatory part of the research was conducted during April – September 2002. The rationale behind using such a period prior to the main study is described in full in Appendix 1. Essentially this work sought to identify the design of a study which was both scientifically rigorous and replicable. Likewise it aimed to inform the selection of measurements and procedures to gauge change in a reliable and valid fashion which was ethically sound. These are all areas, which have been omitted in thinking skills research (DfEE, 1999; McGuinness, 2003).

In terms of prior research a number of criticisms have been levelled at the paucity of research design in controlling for variables adequately (Blagg, 1991; DfEE, 1999; Sternberg & Bhana, 1986). In particular the majority of studies used in effect a *no treatment control* by which to evaluate results (DfEE, 1999). Measures were, likewise, insufficient with no technical detail evident, making replication and scientific study difficult (Bradley, 1983 and Burden, 1987 in Connor, 2002). Moreover, measures focused on a restricted variable of student attainment alone with no triangulation to other domains (DfEE, 1999; McGuinness & Nisbet, 1991; Watkins et al., 2001; Wilson,

2000). The size and scope of samples have also been critiqued as small, with self-selection and optimal teaching environments potentially skewing results (Cotton, 1991; DfEE, 1999; McGuinness, 2003; TES, 2002; Wilson, 2000).

A number of practice questions were left unanswered in the research to date. The key issue of transfer and maintenance of thinking skills was left relatively untouched (Blagg, 1991; DfEE, 1999; McGuinness, 2003; Wilson, 2000). Similarly, the practicalities of the approaches in context with regard to classrooms, staff effectiveness and curricular design were omitted (McGuinness, 2003; TES, 2002; Topping, 2002). Finally the optimum age at which to begin thinking skills instruction was unexamined (Nisbet, 1993) and the duration of time devoted to the thinking skill intervention ignored (Cotton, 1991).

The preparatory phase examined techniques to redress these matters which generate a proposal for action which is original in furthering research in the domain of thinking skills (see Appendix 1 for detail). In such a manner the piece offers one of the first long term evaluations of the infusion methodology, ACTS (Activating Children's Thinking Skills) using both quantitative and qualitative designs. This extends existing research which has examined ACTS at an exploratory qualitative phase only (McGuinness et al., 1997). Earlier restrictions in research will also be overcome through the use of the preparatory phase (see Appendix 1 for more detail).

The research aims to extend work in the sphere of thinking skills by providing a scientific and systematic study of the impact of the ACTS intervention across a cross section of schools. The following sections details how this will be achieved in terms of research design, measures, sample and procedures utilised. A brief critique will then follow to address potential objections.

Research Design

The research is a between groups design in which a range of measures were obtained from two groups of participants from an *experimental* and *waiting list control*. A between groups design was used as a means to compare the effects of the ACTS intervention, by assessing how participants receiving the approach perform, compared to those later receiving to it. *A waiting list control* group was selected as it adequately controlled for the role of expectancy in any given intervention. Furthermore from an ethical perspective it means participants were not denied an intervention, but rather have it delayed. Other control groups were rejected on the basis of limited application to this domain. A *no-treatment control* would not account for the role of expectancy and possible placebo effects. Similarly given the mechanisms and timings of cognitive interventions are unknown, an *attention placebo* would not be credible as the critical aspects are, as yet, unidentified. Finally an *alternative intervention control* would be impossible as the thinking skills interventions to date vary considerably with regards to input, delivery and outcomes. The *waiting list control* group was also considered most ethically sound through consultation with UCL, the LEA and schools involved.

The use of a *waiting list control* is original in thinking skills research and overcomes criticisms of the *no treatment control* traditionally used in this area (DfEE, 1999). The *waiting list control* adequately controls for extraneous variables with the use of randomised allocation to minimise the effects of selection, performance and detection bias (Greenhalgh, 2001).

The study will involve a three year intervention evaluation timescale. Such a period is considered appropriate given the previous research indicating at least two years is necessary in order to generate cognitive change in children (Blagg, 1991; Feuerstein et al., 1980; Shayer & Adey, 1993). The use of a three year period, combined with a *waiting list control* design would allow for exploration into the optimum age (Nisbet, 1993) and intervention duration required (Cotton, 1991) of a thinking skills intervention. Several layers of analysis could be obtained from within group comparisons over the pre, post and delayed post test and between groups over the pre, post and delayed post test follow up. Individual pupil assessment would occur at each of these three points using multi-method measures in a whole class context. In this manner the impact of the intervention over time could be determined through the use of repeated measures. The design would resemble the following as shown in Figure 2 overleaf:-

Time	Year Group Of Pupils	Experimental Group	Waiting List Control Group
Sept/Oct 2002	Year 4	Pre Test Assessment (CATs, MALS, TOPS)	Pre Test Assessment (CATs, MALS, TOPS)
October 2002	Year 4	ACTS training to teachers and intervention implemented for 1 year	No training to teachers and no intervention
October 2002-July 2003	Year 4	**1 year ACTS intervention**	**No ACTS Intervention**
Sept/Oct 2003	Year 5	Post Test Assessment following **1 year of ACTS intervention** (CATs, MALS, TOPS)	Post Test Assessment following **No ACTS intervention** (CATs, MALS, TOPS)
	Year 5	ACTS training to teachers and intervention implemented for 1 year	ACTS training to teachers and intervention implemented for 1 year
October 2003-July 2004	Year 5	**1 year ACTS intervention**	**1 year ACTS intervention**
Sept/Oct 2004	Year 6	Delayed Post Test Assessment following **2 years of ACTS intervention** (CATs, MALS, TOPS)	Delayed Post Test Assessment following **1 year of ACTS intervention** (CATs, MALS, TOPS)
June/July 2005	Year 6	SATs School Assessment	SATs School Assessment

Figure 2; Table to illustrate the quantitative data collection of pupil participants over time to address Hypothesis 1-4.

In this manner the experimental group would receive the ACTS intervention for a two year period (Oct 2002 -July 2004) which could be compared to the waiting list control who would receive it for a one year period (Oct 2003- July

2004).This would address Hypothesis 4 in relation to the optimum duration of an intervention period comparing one year with two years of implementation. The assessment points would focus on Hypotheses 1-3 regarding student outcomes. The use of Cognitive Ability Tests (CATs) would examine the impact on children's thinking (Hypothesis 1). The use of Myself As- A-Learner (MALS) would explore changes in children's self-perceptions of themselves as learners (Hypothesis 2) whilst the use of the Taxonomy of Problematic Social Situations (TOPS) the effect on children's behaviour (Hypothesis 3). Such a research design investigates the independent variable of ACTS on a range of dependent variables across time. The *experimental* and *waiting list control* groups will be comparable in all important aspects except the independent variable being studied. The dependent variables relate to learner outcomes and are detailed in the next section.

Such design of the research ameliorates earlier criticisms of thinking skills study by the provision of an adequate and justified control group (Blagg, 1991; DfEE, 1999; McGuinness, 2003). This extends research to a more scientific examination of a thinking skills approach (Bond, 2002; De Corte, 2002; Gorodetsky et al., 2002). The three year time period, likewise, ensures issues of transfer and maintenance, areas previously ignored in research are explored (DfEE, 1999; McGuinness, 2003; Wilson, 2000). Finally the optimum age and duration of delivery can also be considered (Blagg, 1991; Coles, 1999; Nisbet, 1993). In this way the research design addresses earlier limitations and examines unanswered practice questions in a distinct contribution to the field.

In addition to this quantitative methodology, qualitative aspects are also included to account for earlier criticisms regarding the scope of analysis (Burden & Nichols, 2000; Higgins et al., 2004). These are detailed in a section on Qualitative evaluation and include post-test questionnaires completed by pupil participants and Head teachers alongside pre and post intervention questionnaires completed by class teachers. Such a combination ensures quantitative effects are ascertained alongside qualitative information, allowing insight into the mechanisms at work (Higgins et al., 2004). This is in line with current research in this area (Trickey, 2004; McGuinness, 2003; McKinstery & Topping, 2003).

Quantitative Measures

Measurement selection occurred using the principle of triangulation to ascertain the impact of intervention across a range of outcomes. In particular pre and post measures were administered to gauge the effects on student outcomes including cognitive, behavioural, social and emotional domains. This is in line with current research in this area (see Trickey, 2004; McGuinness, 2003) and addresses earlier research limitations by extending the scope of analysis on students (DfEE, 1999; McGuinness & Nisbet, 1991; Watkins et al., 2001; Wilson, 2000). Furthermore, the use of functional, psychological and social perspectives of outcomes meets criteria of rigorous research methodology (Greenhalgh, 2001).

Two levels of evaluation were utilised namely quantitative measures of students' outcomes and qualitative measures of students, staff and Head teachers involved. In this manner the quantitative data could be built upon and complemented by the qualitative information. Such an approach was adopted to parallel current work in this sphere (Trickey, 2004; McGuinness, 2003) alongside resolving earlier critique of traditional quantitative pre-post designs (Burden & Nichols, 2000; Higgins et al., 2004).

Multi-method Assessment

Multi-method assessment was utilised for students at each of the four points; pre, post and delayed post test of the intervention; Oct 2002, Oct 2003, Oct 2004 and June 2005 respectively. Measures were administered to students on a group level. This was considered in consultation with schools as the most efficient and economical use of time and resources. Individual assessment did not occur as this would be too time consuming and intrusive to students in the school setting. The following measures were administered on the principle of triangulation. As few measures as possible were selected to reduce the impact of spurious results from numerous statistical tests. Technical details of measures used are presented in Appendix 2.

Group Student Measures

Cognitive Abilities Test – Third Edition (CAT 3)

The Cognitive Abilities Tests (CAT 3) was devised by Lohman, Thorndike and Hagen (1993) to assess general reasoning abilities and a pupil's

capacity to apply these to verbal, quantitative and non-verbal cognitive tasks. The device can be used from age 8 to 15 and is appropriate for all abilities. The measure was selected as a method by which to examine student's cognitive thinking skills as specified in Hypothesis 1. The use of the CAT3 was to determine the extent to which children's thinking changes, following the ACTS intervention compared to the control over time. It resolves earlier research limitations by examining students thinking rather than focusing on academic attainment alone (DfEE, 1999). Similarly its use in other thinking skills research at present will allow for comparisons across projects (Trickey, 2004). The measure was also chosen on the basis of its group administration, rather than individual assessment. This was considered alongside the schools as the most economical use of time and least intrusive assessment technique.

The CAT3 is a measure divided into three batteries namely verbal, quantitative and non verbal; to assess general inductive and deductive reasoning skills and cognitive abilities in each domain. Each battery is composed of three separate subtests, thereby producing scores from nine individual subtests, each of the three battery categories and an overall Standard Age Score (SAS). The Cognitive Abilities test will be administered by the two researchers in the classroom context. The students will each receive an answer sheet, question book and scrap paper. The researcher will then explain the activity using a standardised script (see Appendix 3) whilst maintaining general arrangements for testing (see Appendix 4) from NFER-Nelson. The script was designed to alleviate anxiety and stress of

students, whilst also providing a meaningful rationale for occurrence. The children will then work through an example and two practice questions with the researcher. Each subtest will then follow this modelled example lasting between 8-15 mins in total per subtest. A standardised script and activities list was used between each subtest in order to allow for concentration breaks for the children (see Appendix 5). These were devised by researchers in consultation with class teachers in order to minimise the effects of non concentration during the testing procedure. Testing of students in this manner occurred annually to reduce potential familiarity affects (Adey & Shayer, 1994; Blagg, 1991; Feuerstein et al., 1980; McGuinness, 2003).

The CAT3 was chosen as an appropriate measure as it taps into children's thinking and reasoning skills across a range of domains. Likewise, it was considered a more economical and child friendly approach compared with the individual administration of the British Ability Scales (BAS II) or Wechsler Intellectual Scales of Competence (WISC III).

Individual Student Measures

Myself – As –A- Learner Scale (MALS)

The Myself – As – A- Learner Scale (MALS) was compiled by Burden (1998) to examine children's self perceptions of their abilities and approaches to learning. It is a questionnaire which can be completed either individually or with groups of students to provide a measure of children's self concept of themselves as learners and problem solvers. As Frederickson et al (2001)

point out *"the reading and comprehension level required to complete the scale is well within the range of the average 9 to 10 year old"* (p.21). This measure would be read out to the whole class using the script (see Appendix 6 and procedures for detail) to allow for access to all. The questionnaire was considered appropriate for all abilities as it contained items relating to all levels. The measure was selected to ascertain change in students' self-concepts as thinkers, an area hereto untouched in research (DfEE, 1999). It aimed to focus on children's changing self perceptions as learners compared over time, exposure to and duration of intervention as stated in Hypothesis 2. Moreover, its use in similar current studies would allow for direct comparisons across research (Trickey, 2004; McGuinness, 2003).

The measure comprises of a 20 self-referring statements on which each student rates themselves on a 5-point scale. These range from (a) "definitely agree" to (e) "strongly disagree" to allow for positive, negative or neutral responses. The statements were read out to all pupils in a class following a standardised script of introduction (see Appendix 6). This involved an explanation of the measure and purpose, alongside clarifying how to complete. Each statement and five options were then read out to students. This was to ensure that the five negatively worded items were understood and responded to appropriately. Group administration was selected to avoid confusion around such items and ensure all students could fully comprehend the self-referring statements. Furthermore, it meant any reading difficulties would not impair children's understanding or completion of the questionnaire.

Taxonomy of Problematic Social Situations for Children (TOPS)

The Taxonomy of Problematic Social Situations (TOPS) was devised by Dodge, McClaskey and Feldman (1985). It is a questionnaire that teachers complete to identify the specific social situations or tasks a particular pupil finds difficult. It was developed for primary aged pupils. The measure was selected for use as it would provide an insight into children's social and behavioural responses following a thinking skills intervention as stated in Hypothesis 3. Such an area has been largely ignored to date (DfEE, 1999). Furthermore, this measure is currently being utilised in other sister projects which will provide comparisons (Trickey, 2004; McGuinness, 2003).

The measure consists of 44 problematic social situations organised into six types of scenario. Such scenarios encompass peer group entry, response to peer provocation, response to failure, response to success and response to social and teacher expectations. The teacher rates a child on a 5-point scale on the basis of observations and the extent to which the student would experience problems in these given social situations. The 5-point scale ranges from 1, if a given situations is "never" a problem for the student to a 5, if the situation is "almost always" a problem for the pupil. The intermediate points are "rarely", "sometimes" and "usually". In this study the class teacher was given the TOPS to complete for all students in their class. Clear written instructions for completion were provided (see Appendix 7).

Such a device meant a view of both social and behavioural aspects of a student could be ascertained in an economical manner from the individual

who knew the students best. This was considered a more efficient and representative device as opposed to interval or event sampling observations or parental ratings. Observational approaches would not only be more time consuming, but would also require more training to use, with inter-coder reliability a potentially confounding variable (see Dodge et al., 1982 study for detail). Moreover, the measure allowed for an analysis of behaviour across a range of both formal and informal contexts. Finally the Taxonomy of Problematic Social Situations (TOPS) was chosen because it corresponded and tapped into the relationship stated in Hypothesis 3 and is reported to be a measure which is sensitive to intervention effects (Nangle et al., 1994).

Performance and Attainment Measures

Previous research tended to focus exclusively on connecting thinking skills packages to student attainment alone (Wilson, 2000). This study will use such performance measures to ascertain if ACTS, as an intervention, can lead to higher levels in terms of attainment measures as inferred in Hypothesis 4. This focus parallels existing work in the thinking skills domains, in which performance, as measured by conventional national attainment such as GCSEs were used (Blagg, 1991; Shayer & Adey, 1993; Trickey, 2004). For the purposes of this piece the end of academic year 6 will be used focusing on the School Achievement Tasks (SATs).

The rationale for using performance measures in this study is multifaceted. Firstly a focus on attainment ensures this research fits with earlier work in the

area. In addition, it extends study by using both attainment and attitudinal/behavioural measures (Blagg, 1991). The use of SATs also made the research manageable by utilising techniques present in schools. Given the time constraints, this meant intrusive individual testing would not be necessary and schools could conduct assessments independently. The SATs measures were both relevant and workable in the educational environment. This was considered advantageous over other group assessments, such as the NfER-Nelson numeracy or literacy tasks.

Similarly, such assessments are pertinent nationally, which meant findings could be extrapolated to the majority of schools in the UK. Finally, this measure inserted a delayed post test later on in the academic year which had the additional benefit of not being necessarily linked to ACTS and possible expectancy of improvement effects. It also avoided student fatigue in the initial assessment procedure. Likewise it avoided the possible slump in student performance over the summer holidays which may have affected student outcomes in the autumn measurement phases.

Attendance

Participants' attendance rates were varied at the student, staff and school level. In terms of student participants all had parental consent granted on the basis of confidentiality of results. Students were not removed from either the intervention or the assessments on parental preference. Communication between the *experimental* and *waiting list* control was avoided by using

different schools for each group. With regard to staff involved in the implementation of the project in total 26 teachers participated.

At a school level originally 12 schools had been selected to participate in the study. During the start of the intervention, however, 2 schools were removed on the basis of ethical considerations. 1 school was inconsistently implementing the ACTS intervention, alongside an additional thinking skills approach. It was felt that this may contaminate the results and was ethically unsound on participants. The other school was omitted from the project as the teacher who would be implementing the intervention was a newly qualified teacher in a single form entry school. Both the Head teacher and NQT concerned felt the appropriate support would subsequently be unavailable and pressures too great on the individual in the first year of practice.

Process of Quantitative Data Collection

With regard to quantitative data collection it began during the initial planning phase April to September 2002 in terms of obtaining a cross section of Mid-shire schools. Individual participant assessment occurred during October 2002 using multi-method measures in a whole class context (pre-test). This baseline was then repeated one year later in October 2003 to ascertain impact of the intervention on the *experimental* group compared with no intervention on the *waiting list control* (Post test). Such a procedure was used again in October 2004 in order to examine the difference in the

experimental group who had received 2 years of the intervention compared to the 1 year received by the *waiting list* control (Delayed Post Test) tapping into Hypothesis 4. Finally SATs (June 2005) were also used to gauge the long term impact between the 2 and 1 year intervention period of the *experimental* and *waiting list control* group. Qualitative data collection also occurred and is now detailed.

Qualitative Evaluation – Multi-Method Assessment

The research aims to provide a distinctive contribution to the domain of thinking skills research by exploring accounts of those involved in the intervention. Such an approach is novel as it examines the practicalities of a program in a number of real life settings, unlike earlier work in the field (McGuinness, 2003; TES, 2002; Topping, 2002; Wilson, 2000). Additionally the impact on everyday classroom contexts as opposed to optimum educational environments will be ascertained, an area former research neglected (Wilson, 2000). Prior work has been conducted in a relatively isolated manner by focusing on student outcomes alone. This study aims to expand on this by including participants actively involved in the intervention so that a clearer picture of the mechanisms involved are gauged. To such ends qualitative methods will be used to encompass a new perspective into thinking skill interventions in classrooms. It is believed that this combination of quantitative and qualitative data would be more likely to provide relevant information to practitioners in the field.

The qualitative aspect of the research methodology will be complementary to the quantitative, in that it will illuminate areas in an exploratory manner. Earlier research into thinking skills interventions has largely ignored such qualitative data from those involved (Higgins et al., 2004; McGuinness, 1999; Wilson, 2000). Indeed Higgins et al., (2004) highlight in the research there is "*a surprising lack of detail about qualitative data and its analysis*" (p.45). A qualitative focus will, therefore, examine perspectives of the key players. The principle of triangulation will be utilised to obtain student, staff and Head teacher perspectives of the ACTS intervention. These three groups will provide multiple perspectives to examine the effects of ACTS. In particular the changes to class teacher's knowledge and skills development following the intervention will be examined in relation to Hypothesis 5. The pre and post training questionnaire for this group will explore such changes. The use of the Head teacher and pupil questionnaires will serve to investigate these effects as experienced by others. Data relating to Hypotheses 1-4 will also be indirectly examined as to the effect on children's thinking, self-perceptions and behaviours as perceived by participants.

The main method of data gathering will be the use of *semi-structured questionnaires* to explore the mechanisms and impact of the intervention. Such a technique was considered in consultation with schools. *Semi-structured* and *unstructured interviews* were considered too time consuming in the context of the project. Similarly, *focus groups* were omitted given the ethical difficulties of lack of anonymity in the group setting. Schools, likewise, felt this was not an optimum use of student or staff time given the demands

of the curriculum. *Diaries and written accounts* of experiences were discounted on the same grounds. *Participant observation* was also thought to be inappropriate, given the researchers were known to students and staff. Similarly, researcher bias would be likely to play a confounding role in such observations, with subjectivity skewing results to research agendas. *Semi structured questionnaires* met with the constraints set by schools, but would allow an exploratory interpretation into the ACTS intervention in context. This approach was appropriate to both the aims and phenomena being studied.

With regard to data analysis a basic *thematic analysis* will be used given the nature of the information collected. *Discourse Analysis* (Potter & Wetherell, 1987) was considered, however, the use of questionnaires meant naturally occurring language was not being analysed. This meant the data was already being coded and categorised at the point of collection. Such data collection methods meant *Conversation analysis* (Drew, 1995) and *Foucauldian Discourse Analysis* (Willig, 2001) were similarly discounted. Likewise, the focus would not be on the function and constructions of discourse. In this manner discursive psychology methods would not be pure as the study is not *"on how participants use discursive resources and with what effects"* (Willig, 2001) (p.91).

For the purposes of this study it was decided a basic *thematic analysis* would be appropriate as an initial exploration into participant's perceptions, as opposed to a more in-depth theoretically based examination into the nature of discourse. In a similar manner *Narrative Analysis* (Crossley, 2000)

exploring self stories and narrative cognitions were not considered appropriate. *Interpretative phenomenological analysis* (Smith, 1996) with a focus on the person and not the constructions was also considered too individually focused for the purposes of this research. *Thematic analysis* was selected as the means of analysis as it uses discourse *"in its most open sense, following Gilbert and Mulkay (1984) to cover all forms of spoken interaction, formal and informal, and written texts of all kinds"* (Potter & Wetherell, 1987) (p.7). Such an approach would allow for preliminary exploration into evaluative expressions generated through discourse, from the key players involved in the intervention. This aimed to resolve earlier criticisms of the lack of detail regarding qualitative analysis in thinking skills studies to date (Higgins et al., 2004). This procedure is detailed later in Figure 3.

Qualitative Measures

Qualitative data collection used questionnaires as the principal tool to obtain student and staff responses to the ACTS intervention. Such a vehicle was intended to reduce subjectivity in the research process as questionnaires could be completed anonymously by all participants without researchers being present. It was hoped that the absence of researchers during questionnaire administration would reduce the impact on responses.

Questionnaires were devised in consultation with others in the field using a combination of scaled response items and open ended questions. This drew

on work from McGuinness (2003); Trickey (2004); and Watkins et al (2001). No existing questionnaires were appropriate to the purpose required and construction, therefore, built on those currently being used in similar research. This meant technical detail was lacking, reducing the credibility of measures. Items were, however, generated from key issues relating to the intervention, which had been identified in the literature review. Furthermore, consultation with schools also occurred to ensure questions were simple and appropriate. Such discussion meant ethical considerations were examined to ensure potentially sensitive items were removed. Anonymity and confidentiality were guaranteed to obtain an accurate representation of views. In addition, participants were informed that they need not answer questions if they did not wish to. A clear written explanation on the top of each questionnaire explained to participants the rationale in exploring what they thought and a request for honesty. Questionnaires were tested for timing and clarity with three colleagues prior to administration.

Student Questionnaire

The student *questionnaire* was devised to ascertain participant's responses to the ACTS lessons in school. In this manner it aimed to focus on children's perceptions as learners in relation to the intervention as stated in Hypothesis 2. Alongside this it gauged the cognitive gains (Hypothesis 1) and behavioural change (Hypothesis 3) as experienced by the students. The *questionnaire* was considered less intrusive than *semi structured interviews* and less time consuming than *focus groups*. Such a decision was reached in consultation with staff and schools alike.

The *questionnaire* consisted of 9 self-referring questions concerning participant's views about thinking and learning in school. 4 items were ratings on a 10 point scale referring to comparisons between everyday lessons and thinking skills lessons. The key components of such sessions were also to be rated, namely thinking diagrams and thinking groups. Students rated themselves from (0), "Don't enjoy at all" to (10), "Really Enjoy". A neutral was incorporated so as not to force individuals into a particular viewpoint. 5 other items examined what thinking skills children thought they had learnt in the year and the good and bad things about thinking skills lessons. Items also examined both near transfer within, and far transfer outside the school setting. All such issues had been omitted in research to date and offered the first student views on such issues (DfEE, 1999; McGuinness, 2003).

The statements were read out to the pupils in a class, following an explanation of purpose at the top of the questionnaire, by the class teacher (see Appendix 8). This it was hoped would allow for the most honest and accurate participant response. The questionnaire was deemed appropriate for all abilities as it was read out and contained items relating to all. Group administration, likewise, ensured all students could comprehend the self-referring statements appropriately. The questionnaire was administered to all students following the first and second year of the ACTS intervention to gauge response to input (June 2003 and June 2004 respectively). No pre-measure occurred as no point of reference would be available for the students involved.

Staff Questionnaire

Staff involved in implementing and delivering the ACTS interventions completed questionnaires on a pre and post basis in relation to the training phase. This was in order to ascertain the skills and knowledge base prior to and following the implementation of the ACTS intervention tapping into Hypothesis 5. Such an area of research has been omitted in studies to date (DfEE, 1999; McGuinness, 2003; Wilson, 2000). Such an area was explored to gauge the practicalities of the program in real life classroom settings (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000).

The pre-questionnaire was composed of three sections. The first related to personal data to ascertain demographic characteristics of the sample. The second pertained to professional data regarding teaching experience and involvement in thinking skills. A five point scale ranging from (0) meaning "nothing at all" to (5), "a great deal" was then used to gauge skills and knowledge in relation to the teaching of thinking skills. Participants' circled agreement on this scale, a neutral was included so as to allow for the breadth of opinion. A final section obtained data on the class to ascertain how representative they were of the national population (See Appendix 9). This questionnaire was administered at the beginning of the academic year prior to training; September 2002 for the *experimental* and September 2003 for the *waiting list* control. Such a measure aimed to gauge the baseline of process skills and conceptual understanding in relation to thinking skills. The class teacher was asked to complete the questionnaire and return it by post to ensure confidentiality.

This measure was then followed by a post measure questionnaire 10 months later to ascertain any changes, (July 2003 for the *experimental* group and July 2004 for the *waiting list control*). The tool used 4 self-referring statements on the same 5 point scale to quantify differences in the knowledge and skills base with regard to thinking skills (see Appendix 10). In addition a fifth open ended item requested to what extent the ACTS intervention had contributed to professional development. This allowed for analysis into alterations of teacher perceptions in terms of their own practice in the domain of thinking skills. *Focus groups* and *semi-structured interviews* were discounted on ethical grounds and the impact of researcher on responses.

The staff participants also completed a third and final end of year evaluation questionnaire following completion of the training. This occurred in July 2003 with the *experimental* group and July 2004 with the *waiting list control*. This measure was taken from a sister project in Northern Ireland being conducted by Professor McGuinness from Queen's University. The *questionnaire* consists of four main sections, which examine teacher's perceptions of ACTS in schools. In particular the two initial sections focus on views regarding the impact of ACTS on children's learning and on professional development. The third section examines views on ACTS training days and materials. A final section explores future plans in terms of ACTS and teaching thinking generally. The whole questionnaire is composed of 16 open ended response items and a final self referring statement on a 9 point scale to reflect the level of engagement with the approach (see Appendix 11).

The measure was completed on the final training day of ACTS by each class teacher. Clear written instructions for completion were provided on the front of the questionnaire. These were reiterated verbally in order to explain the need for honesty and accuracy in views expressed. Confidentiality and anonymity were assured and time provided for thorough completion. The questionnaires were completed individually so as to avoid any confounds of influencing responses in a group scenario. *Focus groups* were not considered ethical, given the possible impact of LEA researchers and the likelihood of producing socially appropriate participant responses to match expectations. This measure was used as it would provide comparisons across sister projects in Wales and Northern Ireland (McGuinness, 2002).

Head teacher Questionnaire

This questionnaire was constructed in order to obtain an alternative evaluation of the ACTS intervention in classrooms. In particular it was hoped the potential impact of the intervention from an additional perspective in the context of a whole school would be insightful. The questionnaire consisted of 6 open ended questions which related to the ACTS intervention in each school (See Appendix 12). In particular its questions examined the impact of ACTS on individual pupils, classrooms, staff's professional development and the school as a whole. The final two items relate to future plans and provide an opportunity for further comment. In this study Head teachers were requested to complete the questionnaire at the end of the first year of intervention, June/July 2003 for the *experimental* group and June/July 2004 for the *waiting list control*. Written instructions at the top of the questionnaire explained the purpose of the measure and assured confidentiality and

anonymity. A *questionnaire* technique was considered more appropriate as a means of obtaining an accurate and realistic picture of Head teacher perceptions rather than a *semi-structured interview* or *focus group*. Participants completed the questionnaire in a months time scale and returned to the researchers by post.

Procedure

In terms of data analysis the responses were considered using *thematic analysis*. The process of *thematic analysis* was drawn from key articles in the area (Aronson, 1984; Elliott et al., 1999; Leininger, 1985; Taylor & Bogdon, 1984) alongside consultation with Dr. Simon Watt at UCL. These are detailed in Figure 3 overleaf with an exemplar and rationale explained.

Figure 3; Procedure of Thematic Analysis of Qualitative Data

Step	Exemplar	Rationale
1. Data Collection	Questionnaires delivered to 3 groups of participants, namely class teachers, head teachers and pupils. All class teachers completed questionnaires before, during and after the ACTS training. Head teachers and pupils completed questionnaires after 1 year of the ACTS intervention.	The views of participants were considered to establish whether findings were consistent when derived from various groups. It allowed for multiple perspectives on the ACTS intervention to be sought. Class teachers were seen more regularly and were delivering the programme and were therefore followed up over time. This was in keeping with Hypothesis 5 to investigate changes in knowledge and skills. Such a level of analysis was not possible for pupil and head teacher groups given the constraints of time.
2. Transcription	Questionnaire responses were typed up by the two researchers so legible and anonymous for the pupil participant responses on questionnaires. This was a word for word version of responses.	Transcription was necessary for pupil responses given the variability in handwriting. Reponses were typed so as to be legible for coding. Such a step was unnecessary for adult participants where handwriting was easy to read.
3. Coding	Each questionnaire was read and coded by a researcher independently. This involved assigning a code (name) directly onto a segment of text containing that theme or feature. Such a process was guided by considering the transcripts in terms of both "objects" and "subjects" (Parker, 1992, 1994).	Questionnaires were coded by a researcher independently so as not to skew the analysis of information. It ensured a full spectrum of potential codes could be generated. The data was coded alone by two separate researchers so as to reduce reliance on one interpretation alone (Hill et al., 1997).

Step	Exemplar	Rationale
4. Independent Analysis	Each researcher combined and catalogued codes into related patterns and themes. This was defined as "bringing together components or fragments of ideas or experiences" (Leiniger, 1985) (p.60).	Independently researchers identified common codes and collated them into themes. This allowed for the data to be analysed and meaningfully linked together from two different perspectives. Such an approach meant data was not skewed in relation to one researcher perspective alone.
5. Joint Analysis	The two researchers compared and checked codes and themes together from the data. Emerging patterns were identified to encompass the whole data set and range of interpretations. An agreed set of overarching themes and codes were delineated. Data was then recoded. Qualitative variations across themes were examined (Aronson, 1994).	Joint analysis of the data occurred to ensure different conceptualisations were made from both perspectives. This meant a range of possible interpretations were included. These alternative views were combined to agree on a consensus on how best to represent the data with an overarching set of themes. This provided a credibility check to ensure data did not reflect one researcher's opinion alone.
6. Credibility Check - Consensus	The two researchers presented the overall themes to two colleagues to check on construction and consensus (Hill et al., 1997). This involved the two colleagues reading through the questionnaires and evaluating them against the themes generated. Any omissions or misinterpretations were then made.	The inclusion of two colleagues not involved in the research acted as a check on the interpretation and analysis of the data. It ensured researchers had not interpreted the data in relation to their particular perspective given their proximity to the ACTS intervention and research study.

Step	Exemplar	Rationale
7. Credibility Check - Respondent Validation	The researchers presented the themes to the class teacher participants in small group cluster meetings. This was to ascertain appropriate interpretation of the data set (Aronson, 1994.)	The research participants were included in commenting on interpretation of data to add a check on analysis and provide corroboration on themes. It also aimed to allow for further insight in expanding on the researcher's understanding of the ACTS intervention.
8. Credibility Check – Triangulation	Comparisons were made between each set of participants e.g. the class teacher, head teacher and pupil groups by the two researchers. This examined whether key themes were consistent from the range of stakeholders concerning the ACTS approach.	The principle of triangulation was used to establish whether the findings were similar across the data set. This ensured that a range of perspectives on the ACTS intervention were accounted for in the analysis of effects (Barker & Pistrang, 2005).
9. Write Up	The themes were presented in a table for the reader initially followed by a structured account examining each theme in turn. This write up included direct quotations from the questionnaires to ensure the analysis was grounded in the data (Banister, 1994). Interpretations as to the theme were explained with qualitative variations depicted. Inclusion of examples from the data were explicitly used.	Extracts and direct quotations from the questionnaires were used to ensure that the explanations were grounded in the data. In this manner interpretations could be connected to the primary data source of participant responses. This aimed to link the evidence with interpretation clearly. It also allowed the perspectives of participants to be explicitly evidenced.

A number of safeguards were used to protect against subjectivity in this qualitative process. As Potter and Wetherell (1987) point out *"one's own language is constructing a version of the world, while proceeding with analysis of texts and their implications"* (p.182). As a result, as a researcher, *reflexivity* was used as a means to actively consider how experimenter involvement with the matter had determined findings. Firstly the *reflexive process* involved two researchers making explicit to each other and then to two other colleagues the construction of themes. These were then also taken back to participants involved in order to ascertain appropriate interpretation. The actual data was collected in the real life setting of school as opposed to a laboratory. Finally *reflexivity* ensured the researcher's role in the research was scrutinised so meanings were not imposed unchecked. These processes aimed to account for the researcher's contribution to the phenomenon under study.

Attendance

Student attendance for the qualitative data collection at the end of the first year of intervention was varied. This was related to the fact that some students were on holiday as it was close to the summer break. A 100% attendance was obtained from staff involved in the training on both pre and post measures. Head teachers responded with a 100 % response rate on the postal questionnaire. This may be linked to the time of year at administration and the reduced demands towards the end of the year.

Process of Qualitative Data Collection

In terms of qualitative data collection, students were asked to complete a questionnaire following the first year of the ACTS intervention. This meant the *experimental* group completed the questionnaire in July 2003 and the *waiting list control* in July 2004. Staff involved in the ACTS intervention were given three different questionnaires both before and after the training to ascertain changes in skills and knowledge base over a ten month period. This occurred in September 2002 and June 2003 for the *experimental* group and September 2003 and June 2004 for the *waiting list control*. Finally the Head teachers of all schools involved were requested to complete a questionnaire pertaining to ACTS. Such an instrument was used following the first year of the intervention (July 2003 for the *experimental* group and July 2004 for the *waiting list control*).

Sample

Selection of Participating Schools

All primary schools in Midshire were invited to express an interest in participating in the project to develop thinking skills at Key Stage 2. This was circulated in the County's Education Children's Service Bulletin and ensured all schools would be notified and aware. The involvement of Standard Funds finance and the related County Council support may have reinforced initial interest. Schools then had to put forward a statement of interest in participating in the project alongside demographic information pertaining to

the school. Such detail was then used to ensure a representative sample of Midshire schools could be included to provide a meaningful data set. Demographic characteristics were decided upon by the researchers and included the number of statemented students, number of excluded students, number of children receiving free school meals, Year 6 SAT scores and ethnicity statistics. In this manner the research was original in seeking such information in the critical subject sampling phase in the thinking skills research. This attempted to minimise systematic bias to include as wide and diverse a sample of Midshire Schools as possible to be representative of the overall national population.

Schools were selected by two researchers and the Head of Service based on a variety of factors. These included demographic characteristics, consideration of geographical mix of schools alongside a commitment to the release staff to participate in the development, evaluation and training of the approach. Recruitment bias was minimised in two main ways. Firstly, analysis of the School development plan meant schools already enthused and involved in thinking skills packages were omitted as they potentially could contaminate results with alternative programmes. Likewise, although, schools themselves could be skewed in favour of an intervention the staff group to be involved would be designated by the researchers. The schools were unaware at this point that it would be Year 4. Twelve schools were selected on the basis of these factors to ensure a representative sample according to demographic characteristics. This was original in that it included a substantial number of schools from varied geographical locations, which

has been hereto restricted in the thinking skills research (Adey et al., 2002; Blagg, 1991). All schools agreed to participate in the study, except one which felt staff changes had meant the project was impossible to take on. The schools selected were representative of Midshire primary schools and included a stratified and diverse sample of geographical locations and demographic characteristics.

The final selection of schools was then allocated to either the *experimental* or *waiting list control* group. Stratified random allocation of schools to the various treatment conditions was considered imperative to the research design to minimise selection bias. In particular there needed to be no variable, which would erroneously influence conclusions about the groups and distort comparisons. As a result randomised allocation to either group occurred with matched samples according to geographical and demographic characteristics. This ensured the same types of school were represented in both the *experimental* and *waiting list control* group. This meant groups were comparable in all important aspects except the presence or absence of the variable being studied, namely the ACTS intervention. Performance bias was minimised by removing any school that engaged in additional thinking or learning programmes, which could confound results. One such school was removed on this basis. The schools included 4 Primary, 4 Junior and 2 Church Schools.

Participating Pupils

The study involved 404 children, 206 males and 198 females, aged between 7y 6m and 9y 8m (mean age = 8 years 8 months; SD = 4) who were in Year

4 at the projects commencement (September 2002). These were divided into a *waiting list control* and *experimental* group with 244 in the *waiting list control* group and 160 in the *experimental* group. The chronological age for the *experimental* group ranged from a minimum of 7y 11m to a maximum of 9y 8m (mean age = 8y 8m; SD = 4). In the *waiting list control* group the minimum age was 7y 6m and the maximum was 9y 3m (mean = 8y 8m; SD = 4). In terms of the SEN Code of Practice, the sample consisted of 4 statemented students, 27 children at School Action Plus, 48 at School Action and 245 not on the register. These included children with learning, emotional and behavioural difficulties. According to school records the range of languages spoken at home by the pupils included Bengali, English, Punjabi and Urdu. There were no statistically significant differences between the *experimental* and *waiting list control* groups on the basis of gender, age, ethnicity or Code of Practice stage. The individual socio-economic status of individual children was not ascertained, but rather the electoral ward rating of the school which they attended. These were matched across the *experimental* and *waiting list control* group (Greenhalgh, 2001). Detection bias was minimised by measuring the sample with the same materials so no systematic differences in outcomes were made. Instruments were selected to be culturally and ethically appropriate. Furthermore, exclusion bias was reduced as students were followed up at the same point in time over the two year project period.

Students were all in Year 4 for the first year of the project (September 2002 – July 2003) and in Year 5 for the second year (September 2003 – July 2004).

These age groups had been selected so as to parallel research in the sister project in Northern Ireland. Similarly consultations with schools indicated this as the optimum age for intervention in the curriculum. Likewise it would allow for follow up at Year 6 and a two year period of intervention in the same educational environment.

Participating Teachers

Teachers were randomly allocated to the project according to the schools placement in the *experimental* or *control* group. This reduced selection bias as teachers were unknown to researchers and had made no contribution or indication of interest in the projects purpose, thereby, reducing the likelihood of a skewed population. A screening questionnaire was used (See Appendix 9) to ascertain the basic demographic characteristics of the teaching staff to minimise the impact of any confounding variables. In total 26 teachers were involved in delivering the ACTS intervention. This included 8 teachers in the first year of the project (September 2002-July 2003) and 18 in the second year (September 2003-July 2004). There were 16 females and 2 males with a range of years of teaching experience. The range of languages spoken at home included Bengali, Urdu and English. Researchers held an introductory session to the staff explaining the project in more detail and participants were encouraged to share any concerns. Researchers also sought and received permission to use personal data and lesson plans in the course of the project's duration.

Ethical Considerations

Ethical considerations were applied to the treatment of participants according to Elmes et al's 1995 guidelines in Willig (2001) and the BPS (2000). Informed consent about the research procedure was obtained prior to data collection from the schools and staff involved. Head teachers and SENCO's were contacted and meetings occurred in schools to provide an overview of the study. At this point the right to withdraw from participation in the study was given without the fear of being penalised by the LEA. Confidentiality was also assured with all information to be coded and used only in aggregated form to inform future developments of the approach. No individual pupils, teachers or schools were to be identified and the LEA to be disguised according to ethical principles and guidelines (BPS, 2000). Full debriefing of staff and parents involved, concerning the aims of the research was agreed to occur before the beginning of the projects implementation. The use of a *waiting list control* group was discussed with Head teachers as a means of ascertaining differences of intervention duration and a more ethically sound method over a *no treatment control*. Head teachers were made aware of whether they were in the *experimental or waiting list control* before they agreed to participate.

Head teachers then sought informed consent from the school governors and sent out information to the participants' parents via a standardised letter detailing the nature of the project and seeking informed consent (See

Appendix 13). Parents were given the right to withdraw their children from the project without being penalised. Staff participants were, likewise, debriefed as to the research procedure and informed consent obtained before commencing the project. This was conducted by the two researchers so as to ensure staff did have the option to opt out without responses being determined by a School's Senior Management. Confidentiality was assured regarding information about participants acquired during the research process. School staff and parents were guaranteed that information pertaining to the class and children would be kept confidential and returned to the school on completion of the study. All staff were happy to participate in the project and no student was withdrawn on parental request. Debriefing sessions for school staff and the Head teacher were agreed for the end of the academic year and full access to any publications arising from the study guaranteed. The main research piece involved distinctive phases for students, staff and schools as participants. The overview of events is provided in Figure 4 overleaf;

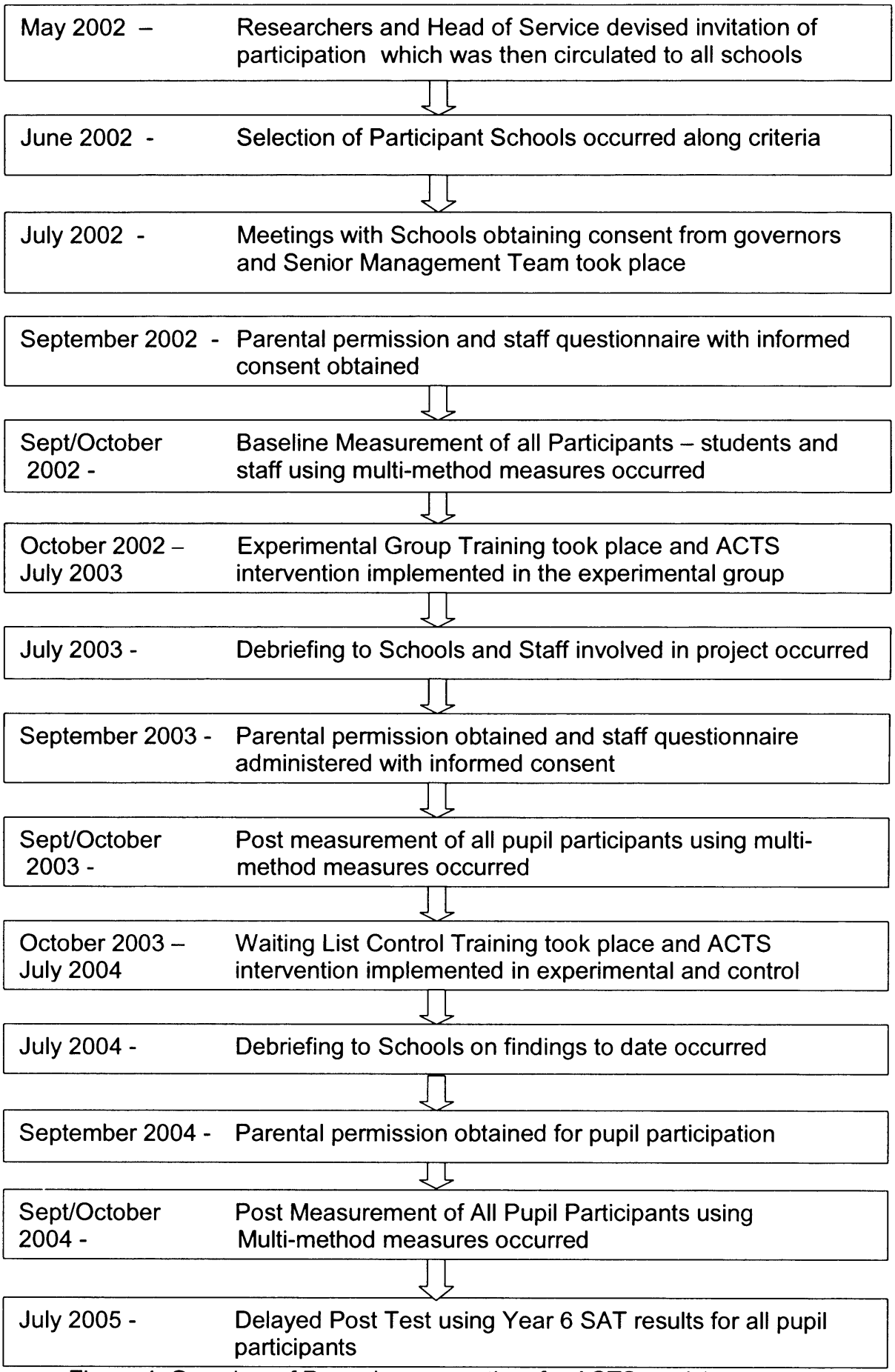


Figure 4; Overview of Procedures over time for ACTS participants.

Procedures for Participating Pupils

In terms of the assessment procedure for pupil participants, this occurred on one day conducted by one of two researchers. The researchers were both female, aged 27 and 29. Researchers swapped classes on the second year of assessment to minimise any experimenter bias. The researchers were introduced to the participants as Educational Psychologists interested in exploring thinking skills. Each researcher then read from a standardised script which introduced the role of an Educational Psychologist, what would happen during the day and the rationale for activities. Confidentiality was also assured and an opportunity to ask questions given (see Appendix 3).

The use of scripts aimed to reduce experimenter bias as did the timing and pacing of the script delivery. This also ensured replication was possible. The assessments then proceeded according to the script in the classroom context. The class teacher was present alongside learning support assistants so as to maintain the normality of the classroom context. Activities were used between sub-tests to allow for breaks in concentration and to reduce the effects of fatigue on students (see script in Appendix 5). Cognitive Abilities Tests (CAT's) occurred first, followed by the Myself-As-A-Learner Scale (Burnett, 1998). Both these instruments have scripts which were adhered to throughout and minimised experimenter bias (see Appendix 3 and 6). General arrangements for testing were also followed throughout according to NfER Nelson guidelines (see Appendix 4).

Procedures for Participating Staff

Staff participants who were involved in the project received a programme of training sessions in connection to which group they were in. Therefore, in Year 1 only the *experimental* group received the training whilst in Year 2 it was the *waiting list control* and *experimental* group. Prior to participation, staff completed a questionnaire designed to ascertain basic demographic characteristics alongside their knowledge and skills base with regard to thinking skills. An explanation and discussion of the project was also provided so as to ensure informed consent.

The training package was based on the ACTS sister project in Northern Ireland conducted by Professor Carol McGuinness. The programme drew on Kolb et al's 1984 adult learning cycle. Initially a 2-day induction session occurred in October conducted by Professor McGuinness. This involved an introduction to the methodology, including modelling infusion lessons, and the thinking skills to be targeted and in what manner. Such sessions acted as the *theoretical learning* and *experiential learning* aspects of Kolb et al's model. Each teacher then went to the next phase of *experimentation with learning*, by returning to their classrooms and teaching ACTS sessions whilst evaluating results. Workshops conducted on half termly basis by the two researchers in two localities then allowed for a *reflection of learning* and practice. In addition 4 full day review days were conducted by Professor McGuinness (February and May) and the two experimenters (April and July). The purpose of these sessions was to reflect on practice to date, provide

feedback and allow for advanced questioning and evaluation of the intervention. The final review day contained a qualitative in-depth questionnaire to provide researchers with staff perceptions of the methodology and to gauge staff development in terms of knowledge and skills (see Appendix 11). Additional evaluations had also occurred in each review day to ascertain training and presentation success. Training, therefore, occurred for 6 training days with the whole group with half-termly workshops in two area localities. This aimed to support staff participants throughout the course of the project whilst allowing insight into the process of implementation. The final review session also allowed for a de-brief of findings to date with the key staff involved to address for researcher reflexivity (Sherrard, 1997; Stevenson & Cooper, 1997).

Chapter 3: Results

Quantitative Data Analysis

Initial Data Exploration

Prior to statistical examination the data set was cleaned to ensure a clear and coherent set of figures. Two main categories were removed, namely SEN status and attainment variables. The Special Educational Need (SEN) status of children was eradicated as there was insufficient representation across all the categories to allow for meaningful statistical comparisons to be made. A similar problem existed with regards to attainment figures as the majority of schools failed to send on the scores relating to the SATs. This meant that hypothesis 4 was unable to tap into the transfer and maintenance of skills as hoped. This aspect had to be omitted from the research due to insufficient data. Finally the NfER Nelson group who scored up the Cognitive Abilities Tests (CATs) lost several classes of data which significantly reduced the original data set. The final set consisted of 118 children in the *experimental intervention* group and 189 in the *waiting list control* group who had full points of data over the 3 assessment periods.

Initial statistical analysis occurred to ascertain how comparable the *experimental* and *waiting list control* groups were. Both the Cognitive Abilities Test (CAT) scores 2002 and the Myself-As-A-Learner (MALS) scores 2002 were normally distributed according to Kolmogorov –Smirnov. The TOPS scores 2002, however, were negatively skewed in both the *experimental* ($K\text{-}S = 0.133$, $df = 139$, $p < .05$) and *waiting list control* group

(K-S = 0.120, df = 225, p <.05). This result is likely to reflect the bimodal distribution found in TOPS scores. Such a finding is to be expected given TOPS scores start at 44 and the majority of children would be scoring at this end of the spectrum. This explains why such a distribution of scores would not be normally distributed.

Preliminary data analysis used two-tailed t tests with a 0.5 rejection level to gauge differences between the *experimental* and *waiting list control* group according to measures and variables. Firstly differential effects have been noted in relation to gender by previous research, such an issue was therefore examined (Adey & Shayer, 1993; Adey et al., 2002). Independent t-tests conducted in relation to gender on all measures found no significant differences on the CAT and MALS scores. However in relation to the Taxonomy of Problematic Social Situations (TOPS) scores and gender, there was a significant difference (t = 3.974, df = 362, p < .05). Such a factor will need consideration in future analyses as it appears boys and girls score differently in terms of their behaviour as rated on the responses to problematic social situations as measured by the TOPS.

The two groups differed with respect to Special Educational Needs (SEN). In the *experimental* group there were 29 children at School Action, 19 children at School Action Plus and 2 statemented children. In comparison there were 19 children at School Action, 8 at School Action Plus and 2 statemented children in the *waiting list* control group. This indicates a far higher number and proportion of children with SEN in the *experimental* group. Overall in the

experimental group 31% of children were on the Code of Practice compared to 20% in the *waiting list control* group. This indicates a factor which could not be investigated further due to small numbers of students representing each of the SEN categories in each of the experimental groups. Subsequently no meaningful comparisons can be made in terms of statistical analysis as insufficient data is available for all SEN categories.

Since the data were normally distributed and had equal variances they met the assumptions of parametric t tests which were conducted on the variables. Likewise the data was obtained using an independent groups design and was an interval level of measurement. An independent t-test was used to compare the means from the two independent groups of individuals i.e. the *experimental* and *waiting list control* group. Overall the analysis suggested that while scores on the MALS and TOPS showed no difference between the two groups there were some differences in relation the CATS. Significant differences were noted between CAT scores according to whether individuals were in the *experimental* or *control* group. There was a significant difference between the conditions in 2002 in relation to the overall scores on the CAT ($t = -4.017$, $df = 368$, $p < .05$). The *experimental* group mean CAT score was 94.67 contrasting to the *control* group's 100.22. This indicates that the *control* group had a higher overall baseline score in terms of the cognitive ability measure. These initial differences will require control in later statistical analysis.

Initial data exploration used box plots to investigate differences between the *experimental* and *control* group. This also allowed outliers to be identified.

Children's Cognitive Ability Development- Hypothesis 1

Children's scores on the Cognitive Abilities Tests (CATs) increased over time as demonstrated in Figure 5. This would be expected given pupil's maturation and cognitive development over the 2 year time period. In addition to this, however, the box plot reveals differences between the *experimental intervention* group and *waiting list control*. In comparative terms the overall standard age scores for CATs increased to a greater degree in the *experimental intervention* group than the *control*. This trend is apparent despite 5 outliers in the control group. These were linked to behavioural issues in the testing context.

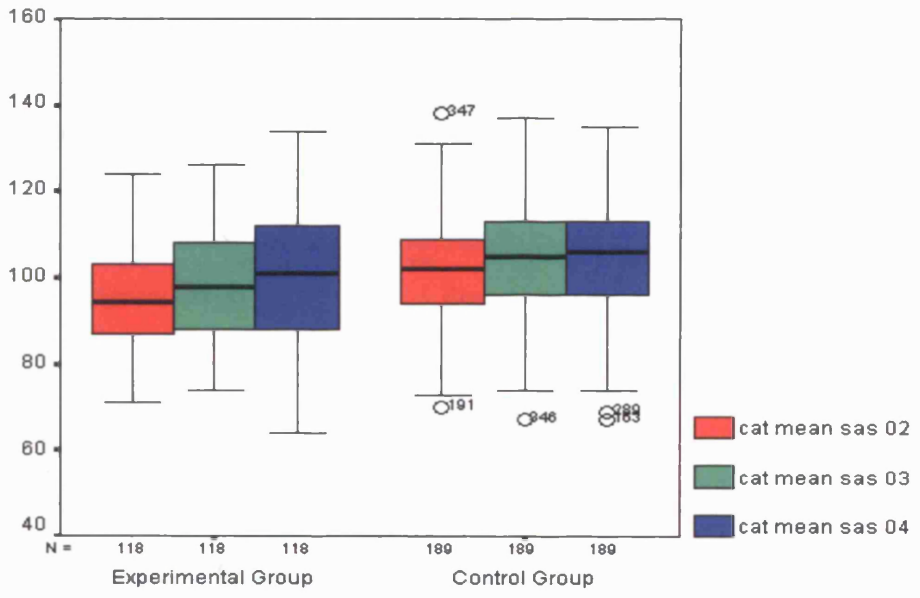


Figure 5; A Box plot to show CAT mean standard age scores over a 3 year period between the experimental intervention and waiting list control groups.

This trend is noted across the subtests of the cognitive ability tests scores. On the verbal subtest, the magnitude of difference on scores over the 3 year period is greater for the *experimental intervention* group over the control. The *waiting list control*, however, made much greater gains initially between 2002 and 2003, which then levelled out in 2004. The experimental intervention group made consistent increases over time. This is clear in Figure 6.

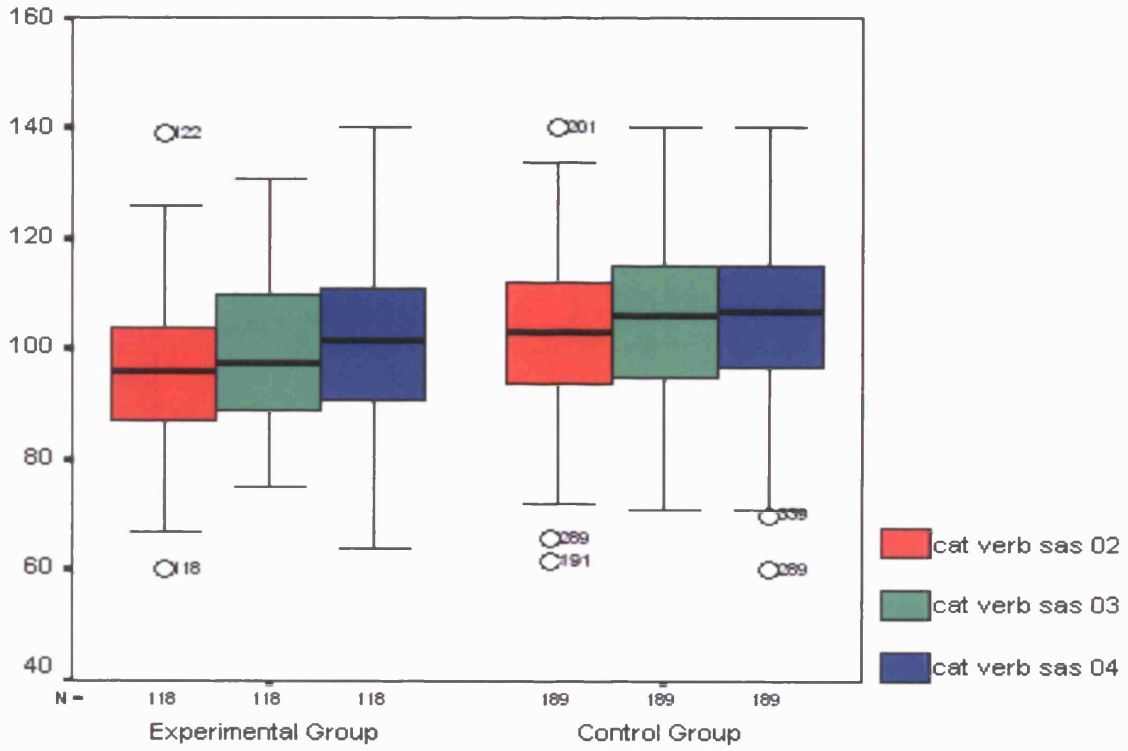


Figure 6: A Box plot to show verbal CAT subtest standard age scores over a 3 year period between the experimental and waiting list control groups.

On the non verbal subtests of the CATs the increase in scores over time is clear for the *experimental intervention* group. The *waiting list control* group, however, shows an increase in the first year but slight decrease in the second year. This is evident in Figure 7. Across both groups the range of scores on the non verbal subtests increases over time.

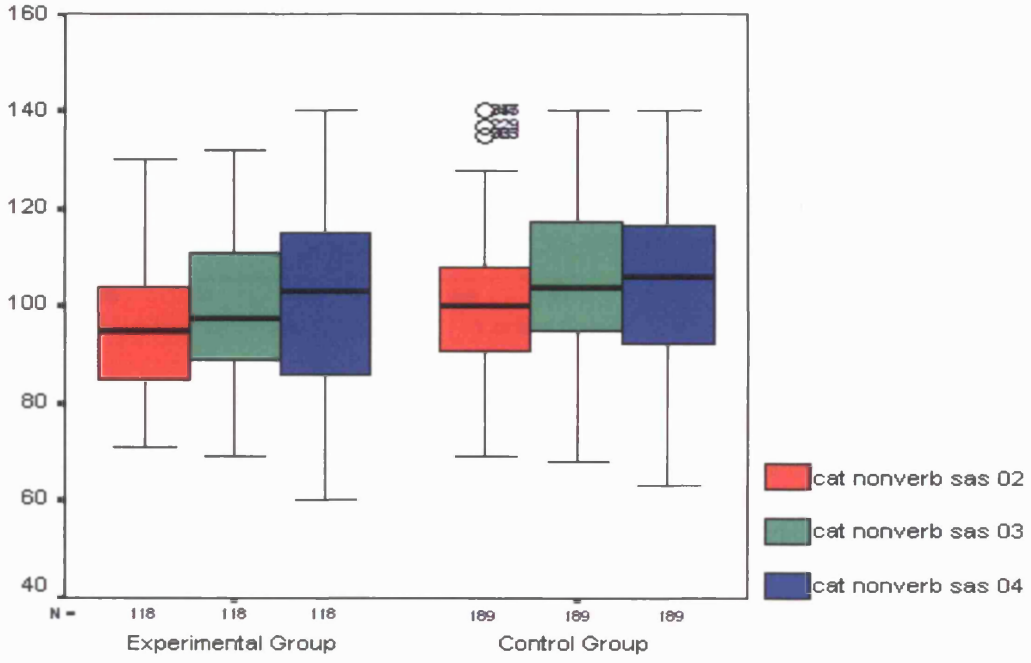


Figure 7: A Box plot to show non verbal CAT subtest standard age scores over a 3 year period between the experimental and waiting list control groups.

Finally the same general score increase is apparent on the quantitative subtests of the CATs as in Figure 8. The *experimental intervention* group demonstrate growth in scores over time which is greatest between 2003 and 2004. The *waiting list control*, however, evidence a more gradual change.

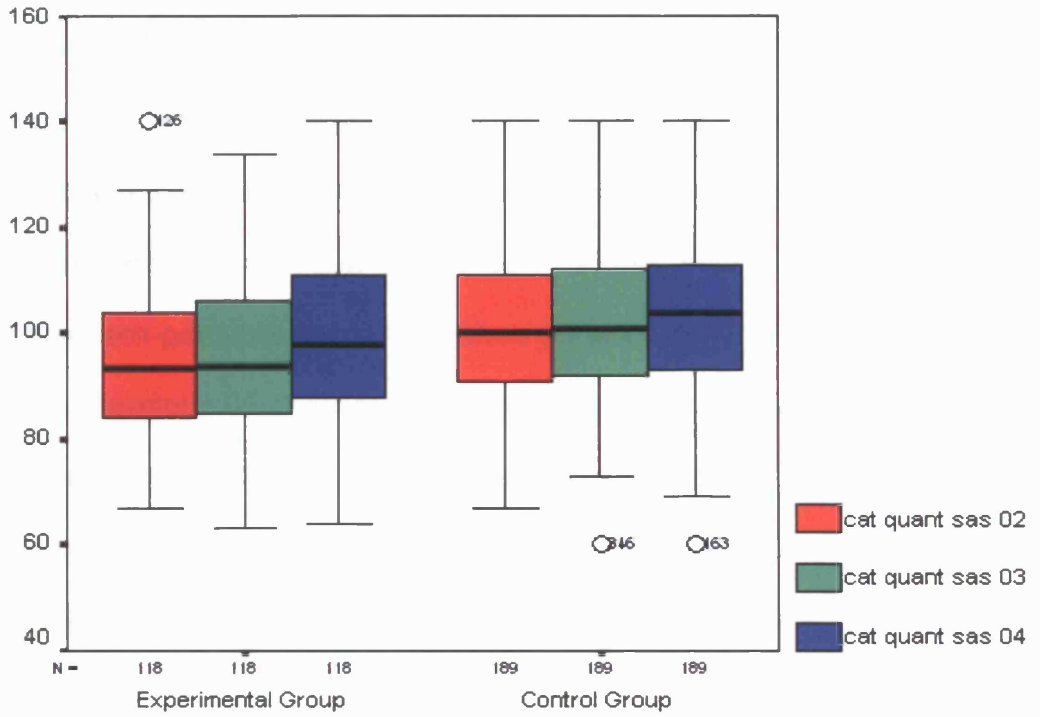


Figure 8; A Box plot to show quantitative CAT subtest standard age scores over a 3 year period between experimental and waiting list control groups.

The initial examination of data suggests a potential impact of the intervention as compared between the *experimental intervention* and *waiting list* control group in relation to Cognitive Ability Test scores as proposed in Hypothesis 1. The initial baseline difference in scores will need accounting for in later analysis.

Children's Self-Perceptions and Behavioural Change- Hypothesis 2 / 3

Children's performance on the Myself-As-A–Learner Scale (MALS) changed erratically over time and between the *intervention/control* groups as shown in Figure 9. In the *experimental intervention* group, scores initially went up for

the first year illustrating pupils' felt more positive about themselves as learners. This decreased, however, in the second year. This may pertain to maturation or more critical self-perceptions of themselves as learners. In comparison the *waiting list control* group scores went down over time across all years. Both groups saw a decline in scores over time reflecting more negative self-perceptions of themselves as learners. This may represent the general developmental decrease in self-perceptions in the elementary years (Burden, 1998; Burnett, 1996). Alternatively it could be that the ACTS intervention made children more aware of what they had yet to know and thus negatively skewed self-perceptions linked to Hypothesis 2.

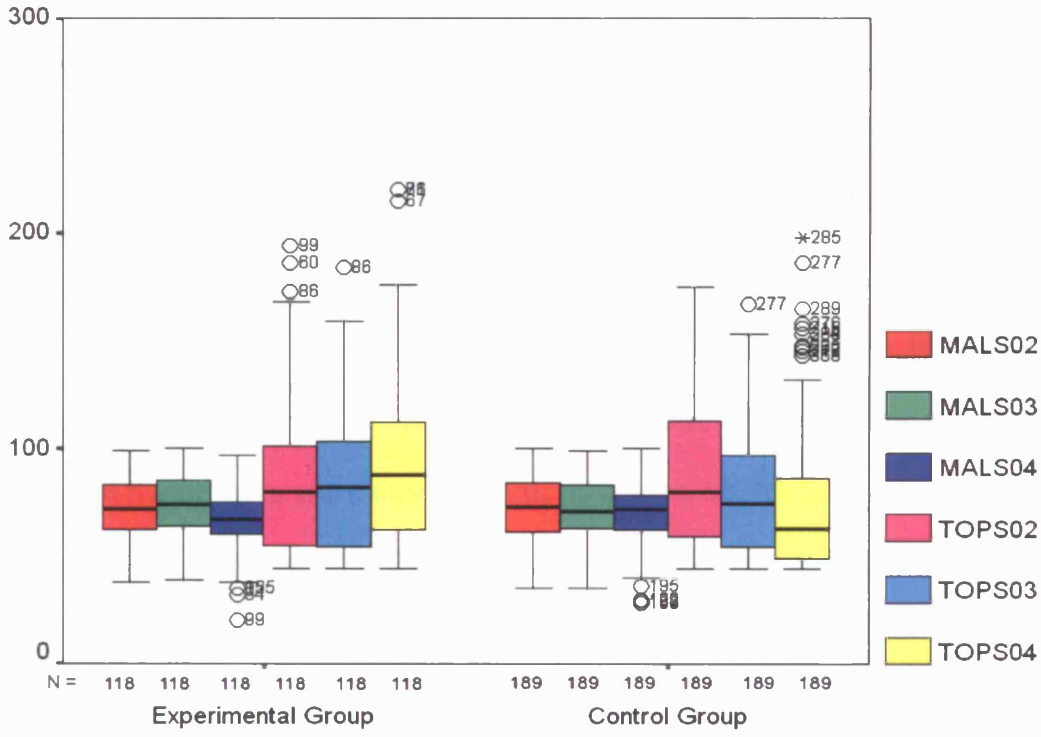


Figure 9; A Box plot to show MALS and TOPS scores over a 3 year period between the experimental and waiting list control groups.

With regards to children's scores on the behavioural measure, the Taxonomy of Problematic Social Situations (TOPS), there is a similarly mixed picture. The TOPS scale measured student's ability to cope with problematic social situations. The minimum score would be 44 and would indicate an individual who is skilled in managing social situations. A maximum score of 220 would indicate an individual who responds inappropriately or ineffectively in most social situations. It was found that over time the *experimental intervention* group obtained increasingly higher scores which marks deterioration in managing behaviour in difficult social situations. In contrast the *waiting list* control group had a decrease in scores revealing an improvement in behaviour in complex social situations. This effect needs further exploration in relation to Hypothesis 3 as it could be attributable to school or class impact. Similarly it may represent a measurement flaw of the device.

Relationships between Measures

Scatter plots and bivariate correlations were examined to evaluate the extent of any correlations between measures. One-tailed t tests were used as directional hypotheses were assumed i.e. as scores on the CAT increased so too would scores on the MALS. This has been evidenced by research in the field (Thorndike & Hagen, 1986).

Children's scores on the Cognitive Abilities Tests and the Myself–As-A–Learner should be linked, in that those who score highly in cognitive reasoning would be more likely to score highly in terms of their self-perception as learners. The converse would also be true with low scores on

the CATS reflecting low scores on the MALS. The correlation between CAT mean scores 2002 and MALS scores 2002 is shown in Figure 10. There was a significant positive correlation ($R=0.377$, $n=370$, $p<0.01$). The MALS (Myself-As-A-Learner Scale) ascertains pupils self-perceptions regarding learning. A minimum score of 20 would suggest an individual who has a very low perception of themselves as learners. In this instance a high CAT score is connected to a high MALS score. Therefore a pupil who had a high self-perception of themselves as a learner would be likely to score highly on the CAT. This has been evidenced by other work in the area (Thorndike & Hagen, 1986). This can be illustrated by Figure 10:

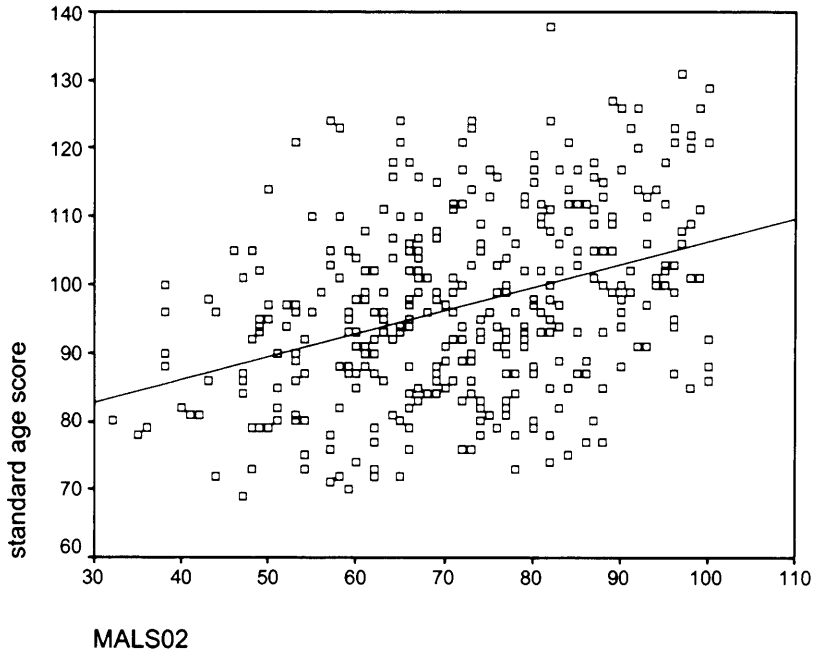


Figure 10; Relationship between Cognitive Ability Test standard age scores 2002 and Myself –As –A –Learner Scale 2002 for the experimental and control groups

The correlation coefficient squared R =0.14. This means that the MALS scores can account for 14% of the variation in CAT mean scores.

Children's cognitive abilities could also be linked to their scores in relation to problematic social situations. The correlation between CAT scores 2002 and TOPS scores 2002 is shown in Figure 11; indicating a significant negative correlation (R = -0.315, n = 370, p = 0.000). It was found that a high CAT score related to a low TOPS score i.e. children who scored more highly on the cognitive measure were more able to manage problematic social situations appropriately.

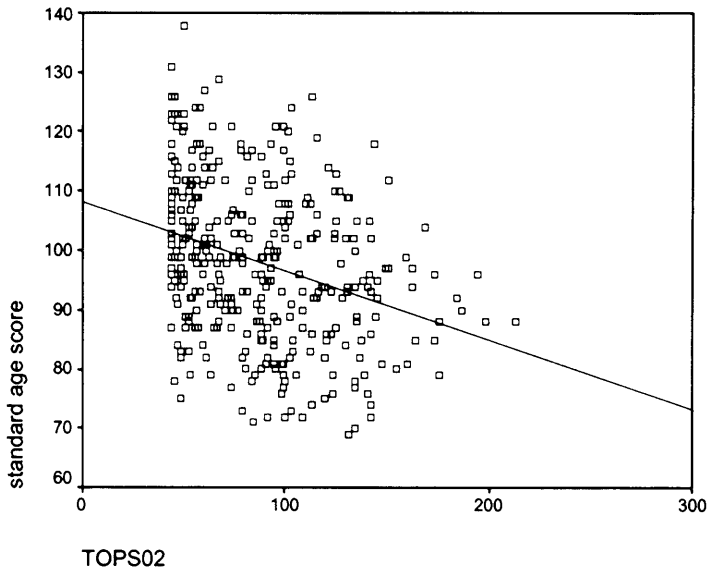


Figure 11; Relationship between Cognitive Ability Scores (standard age scores) 2002 and Taxonomy of Problematic Social Situation scores 2002 for the experimental and waiting list control groups

The correlation coefficient squared R = 0.1. This means that TOPS scores can account for 10% of the variance in CAT scores.

Finally children's' scores on the Myself–As-A-Learner Scale and TOPS were examined. The relationship between MALS 2002 and TOPS 2002 revealed a significant negative correlation as shown in Figure 12 (R = -0.145, n = 364, p = 0.004). This demonstrates that as scores on the MALS increase scores on the TOPS decrease. Therefore an individual who perceived themselves as an effective learner would be more able to appropriately manage social situations as indicated on the TOPS. Conversely a child who regarded themselves as an ineffective learner would be less able to handle problematic social situations.

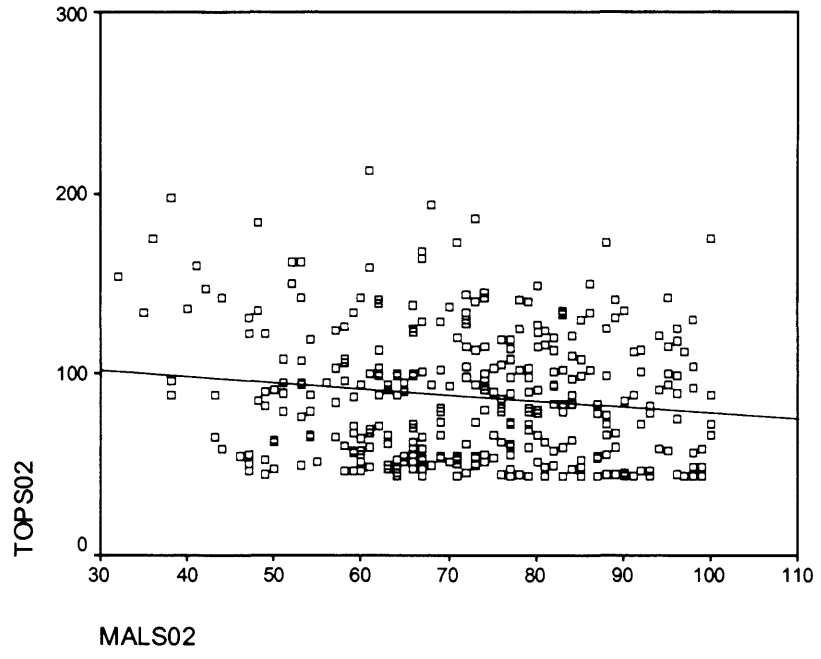


Figure 12; Relationship between Taxonomy Of Problematic Social Situations scores 2002 and Myself- As –A –Learner Scale scores 2002 for the experimental and control groups.

The correlation co-efficient squared R = 0.02. This means that MALS scores can account for 2% of the variance in TOPS scores.

Correlations between measures illustrated a number of relationships, which suggests some interaction between the measures. However this is correlation alone and not causation so could be influenced by some other variables. The scores on different measures utilised indicate significant positive and negative correlations in the direction assumed by the original hypotheses.

Statistical Examination

The ACTS intervention (condition) was examined in relation to its impact on the measures utilised namely the CATs, MALS and TOPS. Multivariate Analysis of Variance (MANOVA's) were used as they allowed experimenters to look at the effect of an independent variable (condition) on a dependent variable (scores on CATS, MALS and TOPS), whilst partialling out the potential effect of other variables i.e. school and teacher. This was a suitable statistical technique to use as subjects were randomly assigned to the condition and the covariates, such as school and teacher, and were measured beforehand. This multivariate analysis allowed one to examine the effect of different independent variables to see if they interact and if there is a relationship between the different dependent variables. The benefits of such an analysis are that within group error variance is reduced and the potential influence of confounds is identified. Significant Multivariate Analysis of Variance (MANOVA) results were followed up with Univariate Analysis of Variance (ANOVA's) or Analysis of Co-variance (ANCOVA's) to explore relationships between the variables further (Brace et al., 2003; Field, 2003).

Children's Cognitive Ability Development- Hypothesis 1

It was predicted in Hypothesis 1 that students would have enhanced thinking skills in the *experimental group* compared to the *control,* this was evidenced in the MANOVA as indicated in Figure 13 in relation to scores on the CATs.

Time	Subtest	F Value	Significance Level	Partial Eta Squared
2002	Verbal	16.541	0.000	0.043
	Non Verbal	8.688	0.003	0.023
	Quantitative	13.605	0.000	0.036
2003	Verbal	12.847	0.000	0.034
	Non Verbal	7.310	0.007	0.020
	Quantitative	11.717	0.001	0.031
2004	Verbal	5.400	0.021	0.018
	Non Verbal	1.871	0.179	0.006
	Quantitative	5.070	0.025	0.016

Figure 13; MANOVA results for Cognitive Ability Tests (CAT) subtests in relation to condition (Experimental/ control group) over time.

This demonstrates significant differences between conditions i.e. the *experimental intervention* and *waiting list control* group on subtests of the cognitive ability scores over time. The only non significant area is on the nonverbal subtest in the 2004 assessment. Despite this the overall picture suggests differences between CAT scores on the basis of group allocation. This does not control, however, for the initial differences at baseline.

Subsequently Analysis of Co-variance (ANCOVA's) were utilised to determine if a significant difference existed between CAT mean scores between the two groups when controlling for pre-test scores. This was achieved using the CAT mean 2002 as a co-variate. Such a decision met assumptions as the co-variate was measured before the intervention began and represented a standardised score. Analysis indicated a between subjects effect to condition which was significant (F= 6.291; P= 0.013, Partial Eta Squared = 0.020). This shows a small effect size pertaining to the condition and is demonstrated by Figure 14.

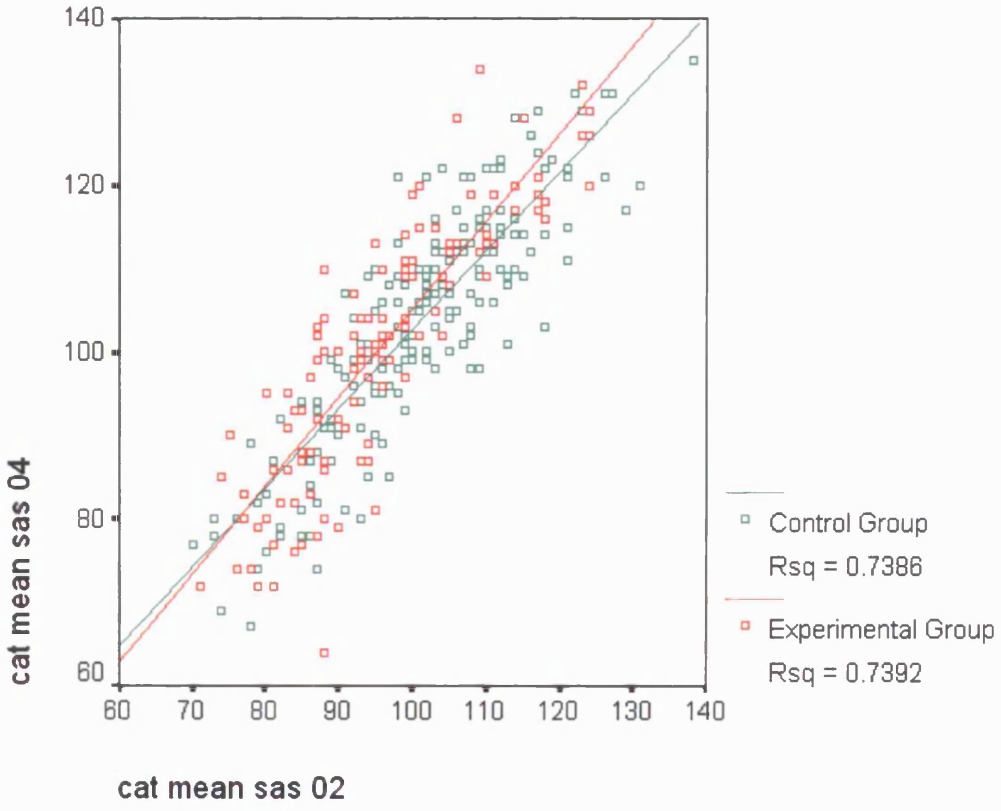


Figure 14; A scatter plot to show the difference between the experimental intervention and waiting list control group in relation to CAT mean scores 2002 and 2004.

This indicates that despite differences at baseline in CAT mean scores between the two groups; the increase in CAT mean scores was greater for the *experimental intervention* group over the *waiting list* control. Such a fact is shown with the CAT mean scores for 2004 being far higher for the *experimental group* compared to the *waiting list control*. To ascertain if this change in CAT mean scores was merely a function of time or the interaction between time and condition a SPANOVA was employed. This enabled exploration into the nature of change between the groups and within groups across the three time periods thus using a mixed between and within subjects design. The following effects were noted as displayed in Figure 15.

Analysis	Wilks Lambda	F Value	Significance Level	Partial Eta Squared
CAT Mean scores	0.784	41.975	0.000	0.216
CAT Mean scores and Condition	0.976	3.695	0.026	0.024

Figure 15; SPANOVA results for Cognitive Ability Test Mean scores and the interaction effect with condition.

The figures illustrate that there was an increase in CAT mean scores over time accounting for 21% of the variance. Similarly there was an interaction effect of condition and CAT mean scores within subjects which accounted for

2.4% of the variance. This represents a small effect size but yet is still statistically significant. Finally between subjects a small effect size is noted for the impact of condition across individual pupils (F= 10.314, p= 0.001, Partial Eta squared = 0.033). This indicates the placement of pupils to either the *experimental or waiting list* control impacted on CAT mean scores as demonstrated in Figure 16.

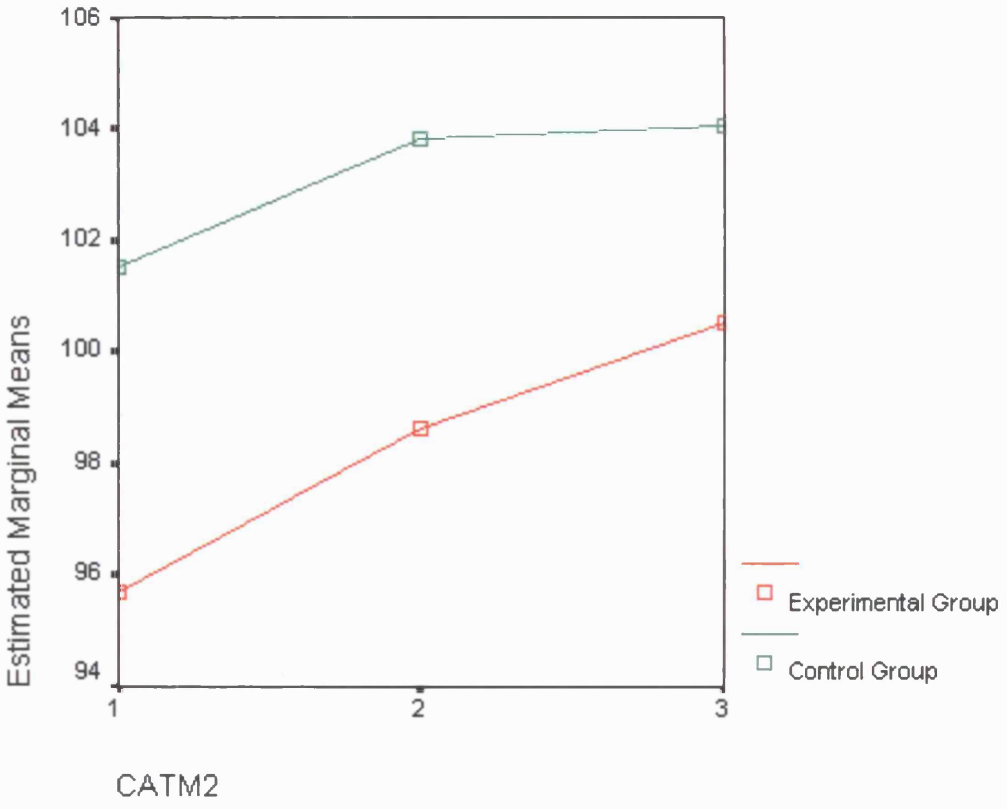


Figure 16: A profile plot to show the difference in marginal means between the experimental intervention and control group in relation to CAT mean scores over the 3 assessment points (point 1-2002, point 2-2003 and point 3-2004 respectively).

As the graph reveals, although, the *waiting list control* group had higher CAT scores initially, these plateaued over time. In contrast the *experimental intervention* group continued to demonstrate increases in CAT scores which were maintained over time. This suggests 2 years of the ACTS intervention is linked to consistent enhanced performance on cognitive ability tests. It supports Hypothesis 1 and 4 in demonstrating improvements over a 2 year as opposed to a 1 year intervention period.

There was a significant effect of Teacher 2003 (8 levels) on the combined dependent variable of CAT scores ($F = 4.030$, $p = 0.019$, Wilks' Lambda = 0.978, Partial Eta Squared = 0.022). Analysis of each individual dependent variable, using a Bonferroni adjusted alpha level of 0.025, showed that there was an impact of Teacher 2003 on the CAT scores 2002 ($F = 7.31$, $p = 0.007$) and on the CAT scores 2003 ($F = 7.887$, $p = 0.005$). Approximately 20% of the variance in CAT scores 2002 is accounted for by the teacher (Partial Eta Squared = 0.020) with 21% of the variance in CAT scores 2003 is accounted for by the teacher (Partial Eta Squared = 0.021). This indicates the effect of teacher on the CAT scores which could generate an effect through a range of different mechanisms. This could be related to age of teacher, years of practice, recency of training and effectiveness of practice. In addition, it may have been that some teachers were more motivated and able to incorporate ACTS into their classroom practice.

Children's Self-Perceptions of Themselves- Hypothesis 2

In connection to Hypothesis 2 it was suggested that children would demonstrate a change in self perception of themselves as learners having received the intervention as measured by the Myself as a Learner Scale (MALS). A MANOVA analysis supported this with a significant effect of Condition (*experimental, waiting list control*) on the combined dependent variable of MALS scores ($F = 5.141$, $p = 0.006$, Wilks' Lambda = 0.970, Partial Eta Squared = 0.030). Analysis of each individual dependent variable, using a Bonferroni adjusted alpha level of 0.025, however, showed that were was no contribution of Condition on individual MALS scores when applied to MALS 2002, MALS 2003 and MALS 2004. It appears that the intervention had no impact on children's self perceptions of themselves as measured by the MALS.

Children's Behavioural Change in Social Situations – Hypothesis 3

Hypothesis 3 stated that children would show a behaviour change in social situations having received the intervention compared to the control group as measured by the Taxonomy of Problematic Social Situations (TOPS). This was indicated by a MANOVA analysis, with a significant effect of Condition (*experimental, control*) on the combined dependent variable of TOPS scores ($F = 9.85$, $p = 0.000$, Wilks' Lambda = 0.944, Partial Eta Squared = 0.056). Analysis of each dependent variable, using a Bonferroni adjusted alpha level of 0.025, showed that there was a contribution of Condition on TOPS 2002 ($F = 13.56$, $p = 0.000$). Approximately 39% of the variance in TOPS scores 2002 can be accounted for by Condition (Partial Eta Squared = 0.039). However

there was no contribution of condition on TOPS 2003 and 2004. This illustrates that the intervention did not affect student's behaviour in social situations at follow up as measured by the TOPS.

There was a significant effect of Schools on the combined dependent variable of TOPS 2002, 2003 and 2004 (F = 8.39, p = 0.000, Wilks' Lambda = 0.952, Partial Eta Squared = 0.048). Analysis of each individual dependent variable, using a Bonferonni adjusted alpha level of 0.025, showed that there was a contribution of the School on TOPS 2002 (F = 7.95, p = 0.005) and TOPS 2003 (F = 15.87, p = 0.000) only. Approximately 23% of the variance in TOPS scores 2002 can be accounted for by Schools (Partial Eta Squared = 0.023) and 46% of the variance in TOPS scores 2003 is attributable to the School (Partial Eta Squared = 0.046). It appears the impact of the school on students behaviour is a more influential factor than that of the intervention.

Qualitative Analysis - Subjective Participant Staff Feedback

The qualitative analysis of end of year teacher evaluations is organised according to key themes identified by thematic analysis of questionnaire responses. These are delineated below in Figure 17.

KEY THEMES	CODES
Theme 1; ACTS as an intervention	1.1 Provides clear framework for lesson delivery 1.2 Explicitly defines different types of thinking 1.3 Infusion across the curriculum 1.4 Time factors
Theme 2; Teacher Development – New teaching strategies and techniques	2.1 Visual diagrams 2.2 Lesson delivery 2.3 Group work 2.4 Displays 2.5 Time to think 2.6 Classroom management

Theme 3; Teacher Development – Language Skills	3.1 Improved teacher questioning 3.2 Teaching language of thinking
Theme 4; Teacher Development – Changes in belief about professional practice	4.1 Shift in attitude towards teaching practice 4.2 Increased self reflection/ evaluation
Theme 5; Pupil Development – Children thinking actively together	5.1 Collaborative work 5.2 Social skills 5.3 Inclusion of all pupils
Theme 6; Pupil Development – Language Skills	6.1 Questioning skills using a thinking vocabulary 6.2 Discussion skills using a thinking vocabulary
Theme 7; Pupil Development – Emotional State	7.1 Enhanced self esteem and emotional intelligence 7.2. Increased confidence
Theme 8; Pupil Development – Application of thinking skills	8.1 Application to problem scenarios 8.2 Transfer to near and far contexts 8.3 Enhanced self-evaluation 8.4 Improved creative thinking
Theme 9; Whole School Development	9.1 Curriculum and planning 9.2 Whole school training

Figure 17 ; Qualitative Analysis of Teacher Questionnaires (2003 and 2004)

The principal themes which emerged from the questionnaire responses are detailed and described in relation to themes and codes accordingly. Reflexivity and validity checks were conducted to ensure plausibility of experimenter interpretation with colleagues and participants alike in their construction. These findings are related to the study's hypotheses more explicitly in the Discussion chapter.

ACTS the intervention as an approach

One theme that emerged from the data was the positive attitude exhibited by teachers to the ACTS intervention, in particular the methodology. They found it helpful that the approach clearly delineated different types of thinking - "*Making me as a teacher think more explicitly about thinking and different types of thinking*" (Susan, Question 1, Theme 1, Code 1.2). This then assisted teachers in identifying classroom contexts in which thinking skills could be applied - "*Aware of different types of thinking that take place in the classroom and when these are most likely to be happening*" (Mark, Question 7, Theme 1, Code 1.2), and in doing so "*ACTS makes the teaching of thinking skills a more accessible vehicle for both the teacher and pupil*" (Amy, Question 1, Theme 1, Code 1.2).

Teachers felt that the ACTS intervention gave them a clear structure by which to organise and plan thinking skills lessons. For example it was reported that it "*provides a clear framework for considering what thinking skills can be used across the curriculum*" (Mary, Question 1, Theme 1, Code 1.1). The easy applicability of the intervention across the curriculum "*encourages the use of thinking skills across the curriculum and to make cross-curricular links*" (Kate, Question 1, Theme 1, Code 1.3). Furthermore, teachers believed this supported the transfer of thinking over subject areas – "*It provides an opportunity to link thinking across the curriculum with user friendly scaffolds*" (Amy, Question 1, Theme 1, Code 1.3). Such evidence was found across the data and was emphasised in the reflexivity checks with participants following interpretation.

Teacher Development- Hypothesis 5

Hypothesis 5 examined the extent to which teachers would have a developed understanding and changed practice to the teaching of thinking skills following the ACTS approach. Several key themes were linked to teacher development in terms of both skills and knowledge in relation to the intervention. These changes impacted on practice and beliefs alike. Firstly, teachers identified a range of new teaching strategies and techniques which they felt enhanced their practice. In particular "*thinking diagrams are useful for structuring thinking*" (Helen, Question 1, Theme 2, Code 2.1) and "*provide clear framework for both teachers and children*" (Chris, Question 1, Theme 2, Code 2.1). These visual organisers had a number of positive effects such as helping "*me to structure lessons and aid the children to think in the most effective way*" (Helen, Questions 7, Theme 2, Code 2.1).

Another practical strategy, which is incorporated throughout the ACTS intervention, is the use of group work. All teachers used groups in the actively thinking together phase of the thinking lesson. This involved all children having a clearly defined role within the group, with the focus on working collaboratively to reach an agreed outcome. Thinking diagrams were used to provide a structure for this process. Teachers felt their skills had developed in the organisation of group work – *"I think I am better at managing/ facilitating children to work in groups*" (Jodie, Question 7, Theme 2, Code 2.3). They could also see the benefits for children as "*they find working in groups reassuring, it helps them to generate more ideas and they enjoy their roles*" (Helen, Question 9, Theme 5, Code 5.1). Teachers also

provided a protected time for children to think as they could see the benefits of doing so – "*I gave the children more time to think and I can see that this is valuable"* (Jodie, Question 7, Theme 2, Code 2.5). This was apparent throughout the data set and supported through validity checks with EP colleagues and teacher participants alike to ensure accurate interpretation.

The ACTS intervention, likewise, impacted on classroom organisation and management. "*The obvious enthusiasm displayed by the kids has infected the general running of the classroom and provided much more social harmony*" (Michelle, Question 7, Theme 2, Code 2.6). This influenced on practice "*because it has forced me to design novel and engaging lessons which the children have enjoyed*" (Amir, Question 8, Theme 2, Code 2.2). The classroom environment was altered with "*displays – engaging children through making them reflect on the thinking process – self assessment reinforcing specific thinking skills*" (Amy, Question 5, Theme 2, Code 2.4). This was evident in a subgroup of a third of participants in the data set.

Teachers reported an increased awareness and improved use of their language in the classroom. This related to "*improved use of a variety of questioning*" (Nadia, Question 1, Theme 3, Code 3.1) and being "*more clear about open ended questioning to stimulate discussion*" (Michelle, Question 5, Theme 3, Code 3.1). "*Teaching the language of thinking*" (Chris, Question 2, Theme 3, Code 3.2) was highlighted as being important and teachers described that they "*use ideas in lessons throughout the curriculum and use vocabulary in questions to develop thinking*" (Mark, Question 13, Theme 3,

Code 3.2). Such findings were consistent throughout the data and highlighted in particular by reflexivity checks with participants.

Thematic analysis revealed that teacher's attitudes and beliefs about professional practice altered. This was firstly in relation to the instruction of thinking – "*I am now a great believer in actually teaching thinking skills whereas before, I thought that developing thinking was a natural developmental stage*" (Chris, Question 7, Theme 4, Code 4.1). This impacted on practitioners' perceptions of themselves in class - "*I see myself less as a giver of facts and decision maker, more as an encourager to enable open discussion and a giver of opportunities for them to make decisions of their own*" (Jackie, Question 7, Theme 4, Code 4.1). This also affected what teachers expected of children in terms of process - "*Before ACTS I was very much focused on directing the children to engage in my thinking and finding my right answers. Now I have realised how important it is not to always spoon-feed the children but let them converse, hypothesise and shape their own thinking when trying to reach a conclusion*" (Amy, Question 7, Theme 4, Code 4.1). Teachers were not as concerned about the need for a written product from every task – "*Not so worried if a child has not written anything down – more aware of their contribution as a thinker*" (Hina, Question 7, Theme 4, Code 4.1). This had a beneficial effect on both children and staff – "*Children are highly motivated – emphasis on talk takes away burden of writing for less able children, which relieves stress on teacher too*" (Mary, Question 8, Theme 4, Code 4.2). Such interpretation was supported by both reflexivity and validity checks.

Finally, there was evidence of increased self-reflection and evaluation on behalf of the practitioners involved - "*I think I am more reflective and self-evaluation of my own practice*" (Mary, Question 7, Theme 4, Code 4.2). In this manner "*it has also helped develop my ability to think about situations in a slightly different way looking at something from all angles and not suffering from tunnel vision*" (Amy, Question 8, Theme 4, Code 4.2). Similarly beneficial effects were noted with a "*more positive and confident outlook on the way thinking can be taught*" (Nadia, Question 7, Theme 4, Code 4.2). This was evident in a subgroup of half the responses in the data.

Pupil Development- Hypotheses 1, 2 and 3

A number of areas of pupil development were evidenced throughout the thematic analysis of data. This included improvements in cognitive, emotional, social and language domains which relate to the areas of development being explored in Hypothesis 1, 2 and 3 respectively. The most predominant feature was that the ACTS intervention encourages children to think actively together with "*children working cooperatively in groups*" (Hina, Question 1, Theme 5, Code 5.1). This meant "*children more aware e.g. taking turns, respecting each other and working using the group rules*" (Michelle, Question 5, Theme 5, Code 5.1). Such joint work had a beneficial impact on social skills as it "*helps children learn from each other and cooperate as a vital life skill*" (Amir, Question 1, Theme 5, Code 5.2). Indeed "*a heightened awareness of social skills amongst pupils*" was noted (Michelle, Question 1, Theme 5, Code 5.2). This connects to Hypothesis 3 in terms of expecting children's behaviour to change in social situations

following intervention. Enhanced social skills in turn also led to improved inclusion in the educational environment where "*all children have learned to value each contribution and have mutual respect for each other*" (Chris, Question 5, Theme 5, Code 5.3). This meant, "*the children all make a contribution and you can see that even SEN children have a valuable contribution to make (often more valuable than other more able children)*" (Jodie, Question 8, Theme 5, Code 5.3).

Another key area of improvement identified was in the domain of language. This paralleled the teacher's development in terms of using thinking vocabulary and enhanced questioning techniques. It was reported, "*children question each other*" (Jodie, Question 5, Theme 6, Code 6.1) and "*are better able to explain their thinking process with a wider vocabulary*" (Mark, Question 5, Theme 6, Code 6.1). This in turn led to greater discussion – "*I feel that the amount of feedback and dialogue I acquired from the children was amazing*" (Nadia, Question 8, Theme 6, Code 6.2). This was particularly noted by teachers working with children with English as an additional language – "*much increased levels of debate amongst EAL children*" (Michelle, Question 1, Theme 6, Code 6.2). This was evident in a subgroup of a quarter of practitioners working with such a pupil population.

Children were reported to make gains in emotional development in terms of both self-esteem and confidence across the data. Comments linked this to the extended vocabulary prevalent in the ACTS classrooms and also to the increased opportunities to use this in small group discussions. As one

participant felt, ACTS lessons "*give children opportunity to express themselves (promotes emotional intelligence)*" (Becky, Question 8, Theme 7, Code 7.1). In particular "*the less able have gained more self esteem and social awareness*" (Michelle, Question 8, Theme 7, Code 7.1). In terms of the development of children's confidence this occurred both due to the fact that "*all children make contributions and are far more aware of their importance in the class* "(Hina, Question 8, Theme 7, Code 7.2) and also "*the fact that less able children have been able to participate more and their confidence has grown*" (Mark, Question 8, Theme 7, Code 7.2). This confidence was linked to the emphasis on structured thinking groups. "*They enjoy their groups and are far more capable. They speak, write, think and interact better – all as a result of their new growing confidence*" (Hina, Question 9, Theme 8, Code 8.1). In particular it was reported that children were "*highly motivated when working in thinking groups*" (Mary, Question 5, Theme 6, Code 6.1).

Finally, children made gains in the cognitive sphere in terms of enhanced thinking skills and their application in a range of contexts. This relates specifically to Hypothesis 1 in examining the development of children's thinking and cognitive development following intervention. It was identified that children "*think in a more structured way*" (Hina, Question 9, Theme 8, Code 8.1). Likewise, it "*improves children's thinking and their ability to tackle different problems*" (Jodie, Question 1, Theme 8, Code 8.1). Teachers noted that children's thinking was not locked in to the classroom context and were beginning to be able to "*transfer new skills to wider world*" (Sue, Question 9,

Theme 8, Code 8.2). Indeed teachers felt they were "*teaching the children skills that they will use throughout their lives rather than facts and figures*" (Laura, Question 8, Theme 8, Code 8.2). In particular children showed increased levels of creative thinking – "*I have been impressed by children's creativity*" (Mary, Question 8, Theme 8, Code 8.4). Again this was linked to changes in group organisation e.g. "*greater creativity, that results from being able to bounce ideas off each other in a non-threatening context*" (Amir, Question 9, Theme 8, Code 8.4). Some metacognitive awareness was evidenced with children "*keen to self-evaluate their own learning*" (Andrea, Question 9, Theme 8, Code 8.2). This was demonstrated in a subgroup of the data set in a third of responses.

Whole School Development

All teachers felt that there would be advantages in establishing the ACTS intervention as a whole school approach. Firstly in terms of planning the curriculum some teachers were "*hoping to match thinking skills objectives with academic objectives on our long term planning before the year starts*" (Sue, Question 13, Theme 9, Code 9.1). It was felt this would encourage teachers to "*use ACTS to a greater extent*" (Michelle, Question 13, Theme 9, Code 9.1). The use of the ACTS intervention was also considered at a whole school level with "*ACTS written into the SDP (school development plan)*" (Ruth, Question 14, Theme 9, Code 9.2). Moreover all schools had plans to develop and incorporate ACTS on a wider scale with "*plans for an INSET day*" (Mark, Question 15, Theme 9, Code 9.2) and staff "*planning to formally introduce staff to ACTS and help staff to begin using thinking skills in their*

classrooms" (Helen, Question 15, Theme 9, Code 9.2). This was consistent throughout the data set and reflexivity checks with participants.

Subjective Participant Feedback: Head teacher Questionnaire

Head teacher's views and perceptions were obtained through the use of an end of year questionnaire. This aimed to establish the effects of the ACTS intervention from a different perspective from within the school. The qualitative analysis and interpretation follows in terms of the impact at a pupil, school and staff level.

Impact on pupils

All head teachers responses indicated a positive impact of ACTS on pupils. In particular the explicit focus on discrete thinking skills was seen as advantageous for children as "*it helps them develop as learners*" (Wendy, Question 1, Theme 8, Code 8.1). It was reported that the approach "*encouraged more logical thinking and an appreciation of the different ways we think*" (Bindi, Question 1, Theme 1, Code 1.2). This links to Hypothesis 1 in enhancing children's cognitive development. A key positive facet of ACTS was considered to be the use of thinking groups. As a consequence of this head teachers felt "*listening skills improved and approach to problems more logical and thought-out*" (Vera, Question 1, Theme 5, Code 5.2). Indeed head teachers felt "*more purposeful group work*" was occurring (Dave, Question 3, Theme 2, Code 2.3). Finally head teachers felt this approach supported children in helping them to "*take risks, be more confident, know it*

is OK to make mistakes" (Wendy, Question 1, Theme 7, Code 7.2). These comments link to the teacher perspective outlined in Themes 1 and 2.

Impact on staff

Head teachers felt the ACTS intervention contributed to teacher's development on a number of levels. This relates to changes in teachers' practice and understanding following ACTS as specified in Hypothesis 5. Firstly classroom management was seen to change with a "*greater focus on organisation – to facilitate delivery*" (Vera, Question 3, Theme 2, Code 2.2). This primarily related to the use of thinking groups. A beneficial impact was also noted in terms of teachers "*incorporating thinking activities into planning*" (Vera, Question 3, Theme 9, Code 9.3) which, was seen as effective. Likewise head teachers reflected that this occurred across the curriculum. Such changes in practice were seen to enhance teacher effectiveness as teachers had "*more experience of managing groups effectively and ways to do this*" (Dave, Question 3, Theme 2, Code 2.3). This was also apparent in the teacher reports (Theme 2). Alternative classroom practice in turn led to a shift in teacher attitudes and beliefs. As one head teacher pointed out "*all staff have moved on in their thinking about how children learn, what it is to be an effective teacher and what it means to be a reflective practitioner*" (Wendy, Question 3, Theme 4, Code 4.2). This was evidenced in practitioners comments also (Theme 4, Code 4.1 and 4.2).

Impact on the School

Responses indicated that all head teachers were keen to develop the ACTS intervention as a whole school approach to the tuition of thinking skills. In particular head teachers described how the approach could be implemented and developed across the key stages using the thinking frameworks and the thinking groups. In connection to this head teachers recognised the need for time to plan and implement such ideas alongside other curricular demands. A substantial benefit of ACTS was that it could be mapped onto the current curriculum; "*It is quite an easy approach to use, as you can use it with existing lessons*" (Dave, Question 6, Theme 9, Code 9.1).

The main limitation seen by teachers pertained to the time needed to plan and implement lessons initially (Theme 1, Code 1.4). This was picked up by only one head teacher who described how there "*were lots of positive responses; staff do feel there is a lot of additional workload but appreciate this burden should decrease with time and experience*" (Vera, Question 6, Theme 1, Code 1.4). Such difficulties could be minimised by a whole school approach incorporating the ACTS intervention into established and existing lessons in a staggered process over time. This would mean staff did not feel bombarded by a new approach but would rather gradually incorporate thinking skills perspectives into the curriculum.

Qualitative Analysis - Subjective Participant Pupil Feedback

The qualitative analysis of end of year pupil evaluations is organised in relation to key themes identified by the thematic analysis examination of questionnaire responses. These are delineated into themes and codes in Figure 18 below. This is then followed by interpretation and discussion.

KEY THEMES	CODES
Theme 1; Awareness of Thinking	1.1 Able to identify thinking skills 1.2 Able to specify lessons in which thinking skills applied (near transfer) 1.3 Able to apply thinking in contexts other than lessons (far transfer)
Theme 2; Collaborative Work	2.1 Paired work 2.2 Group Work
Theme 3; Social Skills	3.1 Listening to other's opinions 3.2 Sharing ideas 3.3 Taking on different perspectives 3.4Supportive and emotional function

Figure 18; Qualitative Analysis of Pupil Questionnaires (2003 and 2004)

Children's awareness of thinking- Hypothesis 1 and 2

In both the *experimental* and *waiting list control* group all children were able to identify the different types of thinking skills they had learnt during the year (Theme 1, Code 1.1). Children could delineate the key thinking skills used by the ACTS intervention. The skills of Problem Solving, Comparing and Contrasting and Decision Making were highlighted as the three most significant. Other thinking skills identified included Sequencing, Creative Thinking and Planning. From these responses it can be seen that children had an awareness of the key thinking skills they were being taught and the processes involved following the intervention. This confirms Hypothesis 1 in that the pupils themselves felt thinking skills had developed following ACTS.

In terms of the application of thinking skills, children were able to specify lessons and areas of the curriculum where thinking skills had been useful (Theme 1, Code 1.2). This ranged from Science – "*Making predictions and creative thinking helps me in Science*" (School H, Class 16, Question 8) to Maths – "*In Maths it has helped me to think about what strategies to use*" (School H, Class 17, Question 8). In more general terms children were able to identify that thinking skills helped to develop overall thinking e.g. "*You can think in a better way and sometimes more logically*" (School F, Class 13, Question 4) and "*The thing I enjoy about thinking skill lessons is that I get to use my brain more than I usually do*" (School B, Class 3, Question 4). This illustrates improvements in self perceptions of themselves as learners as stated in Hypothesis 2. Responses indicated children could see the advantages and rationale in learning about thinking in the school context after the intervention had occurred.

Children were able to apply thinking skills in other contexts outside of school (Theme 1, Code 1.3). This far transfer was linked to academic tasks such as homework – "*They have helped me with my homework and when my sister asks I can explain it to her properly*" (School C, Class 7, Question 9). Likewise, thinking skills were identified as helpful in social scenarios such as stopping fights – "*They have helped me when my brother and sister fight, I use problem solving to help out in that situation*" (School F, Class 13, Question 9). Children were also able to identify very specific situations in which they were able to apply their new skills – "*Which move you are going to do if you play a game like rugby or football*" (School A, Class 2, Question

9). Class teacher comments also reiterated such findings with reports of enhanced problem solving and decision making on school outings and trips. These observations were made by the first cohort of teachers following feedback from parents and other staff. These individuals were not necessarily aware of the ACTS intervention and offer a validity check on findings.

Collaborative Working – Hypothesis 3

Children pinpointed the actively thinking together aspect of the ACTS lessons as a key ingredient of the intervention (Theme 2). This was demonstrated by children labelling both group work and think-pair-share as thinking skills they had learned. Although these are not discrete thinking skills they reveal that children linked the learning of thinking skills to collaborative working. It appeared throughout the data that children picked up on the different way in which thinking was taught in terms of lesson delivery and management. In particular it led to different behaviours in children working together as depicted in Hypothesis 3.

In line with teachers (Theme 2, Code 2.3) children also identified the collaborative group work as a new teaching technique linked to thinking skills. In a similar manner to teachers', children found this beneficial and enjoyable. "*The good thing about thinking skills is that you usually get to work in a group and not come up with ideas on your own*" (School E, Class 10, Question 4). Children connected this new way of working with the enhancement of thinking. For example "*We are thinking and working in a team, ...we are*

learning" (School C, Class 6, Question 4). One aspect of the group work that children focused on was the allocation of roles. This provided children with a clear task to perform. "*I think the good things are that everybody gets a different role so that it's fair*" (School B, Class 4, Question 4"). Groups "*help you to think harder and you can try out all the difficult roles*" (School A, Class 2, Question 4).These responses were evident across the data set.

In addition to this children identified the Think-Pair-Share way of working as another positive aid to learning together. In response to the question "What are the good things about thinking skills lessons?" individuals indicated "*thinking partners and working together in our partners*" (School A, Class 1, Question 4) and that "*its easier with two heads not one*" (School B, Class 4, Question 4). This technique was more apparent in the *experimental* group. Indeed observations by the experimenters of workshops and review days saw this strategy developed by the first cohort of teachers (*experimental intervention* group), to a far greater extent than the second cohort (*waiting list control*). This approach captured the imagination of the first cohort and was, therefore, more widely applied in the context of lessons as indicated by pupil report. This reflects the influence of the group of teachers in each cohort in impacting and influencing each other's classroom practice.

Although children were able to detail the benefits of working together they were also aware of some of the drawbacks. These related principally to social disagreements that occurred in the groups. For example "*Sometimes people start rowing and people don't include themselves so we get less*

ideas" (School G, Class 15, Question 5) and "*Not everyone in your thinking group always cooperates*" (School D, Class 9, Question 5). Issues of people not listening were also raised – "*Everyone doesn't listen to each other*" (School F, Class 11, Question 5). From this data some of the downfalls of group working were noted by a subgroup of a third of the children. These comments, however, were in response to a question which specifically asked about the 'bad things' about thinking lessons. These did not outweigh the positive observations detailed by the pupils regarding the benefits of this way of working.

Social Skills Development- Hypothesis 3

Closely linked to collaborative group work is the development of children's social skills (Theme 3) which connects to Hypothesis 3 in relation to behavioural change. "*I've learnt how to be more creative working in a team and listening to other people*" (School C, Class 6, Question 3). These related to both speaking and listening skills alongside taking on alternative perspectives and emotional inclusion. With regard to listening (Theme 3, Code 3.1) children felt they had made progress in this area and realised the benefits of this. "*I have learnt to listen to other people's ideas as well as my own to make decisions and solve problems* "(School C, Class 7, Question 3). Children recognised that this enabled them to use and combine ideas to generate a variety of options. "*You can put your ideas together to make a very good idea*" (School H, Class 17, Question 4). This opened up opportunities for children to consider alternative perspectives. "*They give you a chance to think about things that you have not had a chance to think of*

before" (School E, Class 10, Question 4). Children were also able to listen to other opinions (Theme 3, Code 3.3), which had a beneficial impact on appreciating the perspective of others. "*Now I listen to other peoples point of view before I make a decision*" (School C, Class 7, Question 9). "*They have helped me by thinking about two sides of a situation"* (School G, Class 16, Question 8). Such responses were evident in over half the pupil population.

With regard to speaking, children felt more able to share ideas in their thinking groups (Theme 3, Code 3.2) – "*You get to share all your ideas*" (School A, Class 1, Question 4). In particular the use of questions as a tool to aid learning were highlighted. "*Asking questions that make you think more about it"* (School C, Class 8, Question 3) because "*questions help you*" (School C, Class 8, Question 3). These types of comments were apparent in a subgroup of a quarter of children.

Children's social skills appeared to be enhanced by the supportive function that the thinking group provided (Theme 3, Code 3.4). Numerous children reported on how the thinking groups helped them to help others. "*I think the good things are that you get to give your ideas to help one another and to help others to understand"* (School D, Class 9, Question 4). Similarly others noted how they received support from their group. "*You get more ideas and different opinions and more support*" (School D, Class 9, Question 4). This supportive function made children feel more included as the groups "*give people a chance"* and *"everyone gets to do it and have a say*" (School F, Class 13, Question 4). This created a culture of acceptance within groups.

"*You can think your own answers first then say your answer then you don't feel stupid about yourself if you get the answer wrong, but as long as you have a go*" (School C, Class 7, Question 4). Moreover children felt able and comfortable to discuss emotions in this supportive setting. "*We get to share ideas and feelings*" (School G, Class 15, Question 4) and "*you can open up and tell what you really think inside*" (School H, Class 17, Question 4). These findings existed throughout the data set. This reiterates Hypothesis 3 in that pupils were aware of changes in behaviour in social situations following ACTS.

Chapter 4: Discussion

Children's Cognitive Development- Hypothesis 1 and 4

Hypothesis 1 stated that *students will have developed thinking skills having received the ACTS intervention compared to the control group.* In addition Hypothesis 4 proposed that *Students will make greater gains following two years of the ACTS intervention compared to one year of the ACTS intervention.* Results from both quantitative and qualitative data analysis show that pupils who participated in the ACTS intervention over a 2 year period made greater cognitive gains than did pupils in the *waiting list control* who experienced the intervention for 1 year alone. Such a time period to generate cognitive change has been evidenced and is in keeping with the prevailing research base (Adey & Shayer, 1993; Blagg, 1991; Feuerstein et al., 1980; Higgins et al., 2004; Lipman et al., 1980; Shayer & Adey, 1993). Triangulated participant feedback from pupils, teachers and head teachers indicated that children had an increased awareness of the discrete thinking skills taught by the ACTS intervention and were able to use these. This is echoed in the wider research field with thinking skills reported to be enhanced (Lipman et al., 1980; McGuinness et al., 1997; Topping & Bryce, 2004).

Furthermore children were able to apply these skills as demonstrated by an increase in cognitive ability scale test scores (CATs) across verbal, non verbal and quantitative domains. This has been shown in relation to Philosophy for Children research also (Trickey, 2004). Similar studies highlight enhanced processing skills and non verbal reasoning of children

following cognitive instruction (Adey & Shayer, 1993; Lipman et al., 1980; Savell et al., 1986). This research is distinctive in illustrating such an impact in relation to an infusion thinking skills intervention. Despite such findings there are a number of issues which require closed scrutiny. These include; the mechanisms of change; the impact of staff and schools; the duration of delivery and finally theoretical implications for learning and cognitive development. These are now examined.

Mechanisms of Change

In determining the critical component to create such cognitive change there are numerous potential candidates. The majority of staff involved in ACTS felt the use of language was crucial in developing thinking skills. All teachers identified the development of a thinking vocabulary as instrumental in enhancing cognitions. Once a shared thinking vocabulary had been developed with children, all staff reported that pupils were able to talk about thinking and articulate the respective processes. This is similar to other research in which the role of language in making thinking transparent and enabling children to verbalise and clarify it was demonstrated (Baird, 1984; DfEE, 1999; Light & Littleton, 1999; McGuinness et al., 1997; Mercer, 1995; Mercer et al., 1999; Trickey, 2004). In this piece teachers also felt the enhanced use of questions were important for pupils and teachers alike. Staff felt that both the variety and use of questions had increased, which meant the processes of thinking were supported. Corresponding findings are evidenced in the communities of enquiry research (Splitter & Sharp, 1995) and Philosophy for Children (Trickey, 2004).

Alternative explanations for developed cognitive abilities could pertain to the use of metacognition in the ACTS intervention. All qualitative feedback triangulated to show that there was an increase in children's awareness of their thinking skills. It may be that this conscious and reflective aspect of ACTS generated cognitive change. Indeed such self-reflection has been identified as crucial (Baird, 1984; Blank, 2000; Cardelle-Elawar, 1995; De Corte, 1990; Doheer et al., 2005; McGuinness, 1990; 1993; McGuinness & Nisbet, 1991; White & Frederiksen, 1998).

Finally the role of social facilitation with children thinking together may be the key catalyst for change. In the ACTS intervention the actively thinking together part of the lesson was highlighted as the most beneficial and influential in generating thinking by both students and staff. It may be the use of collaborative and co-operative thinking activities is conducive to conceptual cognitive change. Such social interaction is recognised in theoretical models of learning, from social construction (Vygotsky, 1978) to cognitive stages of development (Piaget, 1972). Moreover it is indicated as critical in other thinking skills studies (Adey et al., 2002; Higgins et al., 2004; Joiner et al., 2000; Mercer et al., 1999; Wegerif et al., 1999). Such social collaboration and knowledge generation could also be recognised at a whole class level. Indeed powerful learning environments research would suggest all these factors are necessary to move learners from apprentices to experts (De Corte, 1990). These areas will need further exploration in the future to delineate respective roles and impact on generating cognitive change in children.

School and Staff Impact

Although results indicated enhanced thinking skills of the *experimental* group, qualitative variations existed which are indicative of the impact of schools and staff also. These are factors which may have influenced findings, despite not yielding a statistically significant effect. Discrepancies were noted between schools in both student and staff qualitative data. For example in the *waiting list control* group alternative thinking skills interventions were present in pupil feedback such as DeBono's thinking hats. The use of other initiatives alongside the ACTS approach was described as problematic by teachers also in this group in Theme 4, code 4.3. This may have diluted the impact of ACTS through contamination with other programmes. Such a confound may have lead to smaller cognitive gains in the control group and represents a methodological weakness inherent in this research. Differences were also evidenced in terms of far transfer between the two pupil groups. Children in the *experimental* group were far more able than the *control* to detail possible transfer of thinking skills to contexts outside the school setting. This could be attributed to the duration of the ACTS intervention over the 2 year period thus supporting Hypothesis 4 or alternatively the impact of the schools involved. These issues highlight the potential effect of schools on the integrity of intervention implementation.

Discrepancies were found between the *experimental* and *waiting list control* with regards to pupil responses on thinking skills. The *experimental* group were able to list a consistent range of thinking skills whilst the *waiting list control* differed according to the class to which they belonged. For example

some classes labelled thinking skills explicitly whilst others made responses linked to topic and curriculum areas. This finding may suggest that the ACTS intervention was delivered more regularly in some of the *waiting list control* classes than others. Such a suggestion was borne out by the variation in ACTS implementation as evidenced in teacher questionnaires.

Differences were also evidenced in terms of the near transfer of thinking skills. Pupils' responses tended to be linked to the class to which they belonged. This once again may represent individual teachers' preferences and skills in applying ACTS. Such findings illustrate the potential power of teacher impact on the effectiveness of an intervention as evidenced elsewhere (Adey et al., 2002; Blagg, 1991; Higgins et al., 2004; McKinstery & Topping, 2003).

Optimum Duration of Delivery- Hypothesis 4

Hypothesis 4 stated that *Students will make greater gains following two years of the ACTS intervention compared to one year of the ACTS intervention.* The study demonstrates cognitive change within the junior age range following a two year delivery of the ACTS approach. It confirms a two year duration as the minimum to create cognitive change (Blagg, 1991; Feuerstein et al., 1980). Indeed it is consistent with other programmes in evidencing cognitive gains only following at least a two year period (Adey & Shayer, 1993; Lipman et al., 1980). This piece illustrates that two years are required to generate significant cognitive change as measured by CATs in the primary age range over and above a one year duration period.

More recent research, however, has produced more immediate effects in a shorter time period from year 1 pupils following a years intervention (Adey et al., 2002) to those in secondary schools following a termly intervention (McKinstery & Topping, 2003). This has important implications to the theoretical debate as to both optimum age and duration delivery (Coles, 1993; DfEE, 1999; Nisbet, 1993). Certain thinking skills interventions may generate cognitive change at quicker rates according to the child's age. For example the paired thinking skills approach (McKinstery & Topping, 2003) may yield speedier cognitive gains linked to the fact the children are cognitively more developed. Alternatively it could be that children are amenable to certain forms of cognitive instruction within curriculum contexts. Such an area is still in need of further examination and exploration (DfEE, 1999). This is particularly pertinent given the implications for curriculum design and pedagogy within thinking schools and potential application in educational practice (TES, 2002; Watkins et al., 2001).

Learning and Development Distinctions

An avenue for discussion surrounding the nature of cognitions is opened up by this research. In this piece children's cognitive abilities were enhanced to a greater extent following 2 years of the ACTS intervention. The question remains, however, as to what was actually changed; for example did children merely learn about thinking skills or did some cognitive development at a deeper level occur. Adey et al (2002) would argue interventions such as cognitive acceleration are indicative of cognitive development, as to generating irreversible progression linked to cognitive growth as opposed to

learning alone (Inhelder & Piaget, 1958). Such suppositions are less easy to prove in the context of ACTS given no piagetian reasoning tasks were utilised to highlight such effects.

In this research there is evidence of transfer, however, which suggests generalisation of learning at a higher level. Qualitative triangulated data from students and staff illustrate thinking skills transferred to both near and far contexts in the *experimental* group. This indicates cognitions are transferred beyond the context of acquisition (Salomon & Perkins, 1989). In future the durability and longevity of these findings would be required to determine if indeed cognitive development has occurred (Adey et al., 2002; Coles, 1993; DfEE, 1999; Georghiades, 2000). Similarly more rigorous research tools to ascertain the impact of cognitive development and learning would assist in differentiating between the two (McKinstery & Topping, 2003).

Children's Self-Perceptions of Themselves- Hypothesis 2

Hypothesis 2 stated that *Students will demonstrate a change in self-perception of themselves as learners having received the ACTS intervention compared to the control group.* In the quantitative data statistical analysis did not evidence a significant effect of the approach on pupils. In fact all that was noted was a slight decline in scores over time for both groups on the Myself-As-A-Learner scale. This could be explained by children's maturation, for as they become older they may become more critical of themselves. Such a developmental trend is apparent in research with a decrease in self-concept in the elementary years (Burden, 1998; Burnett, 1996).

Additionally it may be that the measures utilised in this study were not sensitive enough to detect change in self-perceptions. Such a possibility is substantiated by the qualitative participant feedback. Staff involved commented on the increased self-esteem of students following intervention and general enhanced confidence levels. Such effects mirror research across the thinking skills sphere (Fisher, 1999; McKinstery & Topping, 2003). Students involved in ACTS, likewise, described how they felt more comfortable and confident in volunteering ideas. Such pupil report is also evidenced in Philosophy for Children research (Trickey, 2004). These represent, however, perceived gains alone and illustrate a range of points to ponder. These form the focus for the following discussion namely the variables for impacting on pupil self perception alongside practice and research implications. The qualitative data suggests some perceived gains for children in the emotional domain after ACTS in line with Hypothesis 2.

Variables of Impact

Teachers involved in ACTS reported a positive impact on children's confidence, self-esteem and emotional intelligence (Teacher Theme 7, Code 7.1). Similar findings were made by staff in the Philosophy for Children initiative in Clakmannanshire (Trickey, 2004). This could be expectancy of improvement on behalf of the staff rather than actual pupil gains. Other research, however, links such changes to children' self-perceptions to positive teacher attitudes (Blagg, 1991; Topping & Bryce, 2004). In some ways the enhanced staff belief in pupil's thinking could generate a self-fulfilling prophecy effect in which children who are expected to improve

therefore do so. Indeed the impact of teacher's statements on student's self-concepts has been evidenced (Burnett, 1999). Alternatively it could be argued that encouraging children to think creatively with no one right answer increases pupil confidence and self-esteem (Biggs & Moore, 1993; Burnett & Proctor, 2002).

The fact that the ACTS intervention encourages a culture of enquiry and acceptance of all view points may mean children feel more confident to contribute. This is evidenced in pupil feedback in which children reported being able to discuss matters more openly. Children detailed how everyone had a chance and were able to participate in ACTS lessons. Similar statements were made by pupils and teachers in the Philosophy for Children research alongside video evidence (Trickey, 2004). More open and non-judgemental dialogue was noted. The lack of emphasis on written outcomes may also have added to children's self-esteem and improved self-perceptions as learners. Such a possibility was identified in student and staff comments. It appears that a positive impact indicated in the qualitative data requires further examination with the mechanism for change explored.

Implications for Practice and Research

With regards to the reported enhancement of self-esteem and self-perceptions in pupils, teachers also felt an impact was made on the classroom ethos. In particular staff felt the use of collaborative group work and whole class discussions into thinking, increased the participation of all children in ACTS lessons. In this manner the intervention supported inclusion

of all learners, including children with special educational needs and English as an additional language (Theme 5, Code 5.3). Qualitative staff feedback indicated that these particular groups of children had grown in confidence and were far more able to participate given the structure and support of the ACTS lessons. Evidence in other research supports that there may be a greater impact on low attaining pupils (Cardelle-Ellawar, 1992; 1995; De Corte et al., 2001.) The implications of this for classroom practice in generating an open and flexible forum for thinking and learning needs pursuing. Similar findings have been made in Australia with teachers more skilled in teaching children with learning difficulties using thinking skills approaches (Stewart & Smardon, 2002). In addition philosophy for children and communities of enquiry research has yielded confirmatory findings (Splitter & Sharp, 1995). Future work will need to examine the impact of thinking skills packages for different cultures (Dart et al., 2000a) and varying special educational need groups (McKinstery & Topping, 2003).

The impact and importance of thinking skills interventions on children's self-perceptions and self-esteem requires substantial scrutiny. Indeed considerably more systematic study is necessitated with sensitive measurements to quantify change (Burnett & Proctor, 2002). Moreover, the connections between children's self-esteem and self-concept as a learner need exploration in relation to the learning environment (Dart et al., 2000b). Dispositions in relation to thinking and learning, likewise, require closer consideration (Nisbet, 1993; Perkins et al., 1993). These are detailed in more depth in later sections.

Children's Behavioural Change in Social Situations- Hypothesis 3

Hypothesis 3 stated that *Students will show a behavioural change in social situations having received the ACTS intervention compared to the control group*. Quantitative results using multivariate analysis indicate a non significant impact of the intervention on the Taxonomy of Problematic Social Situations scores (TOPS). It appears that the intervention made no impact on children's behaviour in problematic social situations as measured by the TOPS. Instead the key factor affecting these scores was the school to which pupils belonged. This mirrors the majority of research in this domain (DES, 1989; Galloway & Rogers, 1994; Jenner & Gravenstede, 1998). It may be that the TOPS measure was too focused on problematic behaviours alone as the qualitative data suggested several positive effects in relation to student's social skills development. Perhaps in future focusing on more general social skills would be more appropriate and adept at detailing respective change.

In the qualitative data teachers and pupils identified improvements in a number of social skills therefore indicating behavioural change in relation to Hypothesis 3. These abilities firstly relate to enhanced speaking and listening skills. Children were able to reflect that they were able to listen more, whilst staff detailed how children were more aware of taking turns, listening to and respecting others views (Teacher Theme 5, Code 5.1). Pupils felt far more willing to contribute and questioning was identified as being extended (Teacher Theme 6, Code 6.1). Moreover, children felt they were more able to work in teams and staff believed that working co-operatively in groups had developed (Teacher Theme 5, Code 5.2). Both these effects were noted in

Clackmannanshire research into Philosophy for Children (Trickey, 2004). It would seem that a number of social effects were made but debate persists as to the mechanism of change.

One hypothesis could be that the impact of the intervention was positive in developing children's ability to manage and cope appropriately with social situations. This information, considered alongside qualitative analysis, could be attributed to the emphasis on thinking actively together in co-operative groups. Significant numbers of staff highlighted the impact of group work on the ability of children to listen, problem solve and take on the perspective of others. This links into other research which suggests that high quality social interaction can affect cognitive functioning (Joiner et al., 2000; Light & Littleton, 1999; Ritchie & Edwards, 1996; Wegerif et al., 1999). It could be that this represents an additional positive effect on social skills development. It could be that the use of thinking groups or actively thinking together aspects of ACTS lessons enhances interpersonal skills. This could be linked to the time given to develop these skills or the clearly defined roles in thinking groups. These effects could, alternatively, be accounted for by children's maturation and development in social skills linked to age and experience.

The key variable evidenced as generating an effect of statistical significance was the school. The behavioural policy and methods for tackling problematic social situations in different schools may have determined responses on the TOPS scores. This would need further exploration in the future. At present the impact of the school exerts more influence on pupil's behaviour than the

ACTS intervention. This is concurrent with present research (DES, 1989; Galloway & Rogers, 1994; Jenner & Gravenstede, 1998).

Teachers' Professional Development- Hypothesis 5

Hypothesis 5 stated that *Teachers will have a developed understanding and changed practice to the teaching of thinking skills following the implementation of the ACTS intervention.* Qualitative data evidences enhanced teacher effectiveness across a range of domains from new and alternative classroom practice to changes in professional beliefs. Such findings resonate with research in the field (Blagg, 1991; McGuinness et al., 1997; McKinstery & Topping, 2003; Munro, 1999; Stewart & Smardon, 2002; Topping & Bryce, 2004). Despite such a positive effect, however, discrepancies in the data exist which may explain the differential impact between the *experimental* and *waiting list control* group in relation to children's cognitive gains. These relay to the integrity of the intervention implementation which is critical in determining effectiveness (Adey et al., 2002; McKinstery & Topping, 2003). Each of these areas of development and discrepancy are now discussed.

Infusion as a technique to teach thinking

The qualitative data highlighted that teachers felt that the principle of infusion was key in applying the teaching of thinking across the curriculum. Both Theme 1 and Theme 10 indicated that the ACTS intervention provided teachers with a clear framework for the inclusion of thinking skills into lesson planning and delivery. The explicit nature and structure of the approach was

identified as assisting this process. An increased awareness of the different types of thinking was noticed across all staff. This in turn meant staff placed a greater value on the tuition of thinking skills within the existing curriculum. Such feedback is in keeping with other research in this area (Higgins et al., 2004; McGuinness et al., 1997).

There were some inconsistencies, however, amongst the data set with regard to the implementation of infusion within the intervention. Whilst all participants were able to deliver lessons using the principle of infusion there was variation between Cohort 1, the *experimental* group and Cohort 2, the *waiting list control*. Teachers in Cohort 1 planned and delivered more discrete thinking skills lessons in all areas of the curriculum. In comparison staff in Cohort 2 were more reluctant to plan discrete thinking skills lessons as they saw this as being "*rather regimented*" (Jackie, Question 2, Theme 1, Code 1.3). Instead this group used key elements of ACTS intervention lessons alongside other thinking skills initiatives already in place in their schools e.g. Philosophy for Children and De Bono's six Thinking Hats.

Qualitative variations existed between schools in terms of how the ACTS approach was used alongside other initiatives. This was more evident in Cohort 2 with teachers noting the difficulties in "*fitting in with other initiatives*" (Laura, Question 2, Theme 4, Code 4.3) with an "*overload of strategies e.g. de Bono, CORT*" (Becky, Question 2, Theme 4, Code 4.3). These variations were linked to within-school factors and qualitative comments from the *waiting list control* group suggest conflict surrounding the range of thinking

skills interventions being used. Potentially this might have lead to the dilution of approaches and subsequently a reduction in their effects. One solution would be for schools to adopt a uniform and consistent approach to the teaching of thinking.

New and alternative classroom practice

An underlying consistency across the data was the use of new and alternative teaching techniques and strategies, which embedded the teaching of thinking within the curriculum. All staff felt that their knowledge base and understanding of thinking skills had been extended. This is evident in evidence from Head teachers also and research (McKinstery & Topping, 2003; Topping & Bryce, 2004). In addition to this, participants' skills in applying this knowledge to the classroom had developed through the use of visual diagrams, collaborative group work and a shared thinking vocabulary. This was evidenced across themes 1, 2 and 4. This supports suppositions in Hypothesis 5. Similar findings emerged in the McGuinness et al. (1997) study in which the development of teachers' knowledge and skills were reported.

Both groups of staff used techniques for "*actively thinking together*" across the curriculum and could see a range of benefits linked to this. This was incorporated in the second part of ACTS intervention lessons and involved collaborative group work (thinking groups) and thinking diagrams. The use of thinking groups was the strategy most widely applied by all staff in a range of curriculum areas. This involved each member of a group having a distinct role (e.g. scribe, thinker, chair, listener, timer). Staff highlighted this had a

positive impact on classroom organisation. It led to improved inclusion, where the contributions of all children were valued. Improvements in teaching techniques have been apparent in similar studies (Stewart & Smardon, 2002).

Although all staff utilised thinking groups there were variations in the manner in which groups and roles were defined over time. In the *experimental* group, the roles and the members of the thinking skills groups were consistently maintained over time. This was in order to allow children to acclimatise to the different way of working. In the *waiting list control*, groups were used more flexibly with roles and groups changing over time. This may have been linked to the chronological age of the children and their ability to interact with others in a joint activity. Similarly some of the children in the *experimental* group were experiencing this way of working for the second year and so were more comfortable with these thinking groups.

There was also qualitative variation within the data set in the way thinking diagrams were utilised. Whilst Cohort 1, the *experimental* group, found them a useful structure to guide the thinking process, Cohort 2, the *waiting list control*, at times found them restrictive and felt that they wanted to "*challenge the thinking beyond the diagrams*" (Andrea, Question 2, Theme 2, Code 2.1). The use of thinking diagrams as a scaffold to thinking may have been more applicable and appropriate to the age of children in Year 4 with Cohort 1. Alternatively the use of these techniques may have been influenced by individual practitioner preference. In Cohort 2 certain viewpoints were

strongly indicated during review days and workshop sessions. This may have impacted on Cohort 2's beliefs and practice. Despite this variation, evidence indicated that all staff appreciated the need to teach thinking in all areas of the curriculum. Furthermore they were able to use various aspects of the ACTS intervention to achieve this.

Professional Development

The inclusion of new classroom practice led to teachers making professional development gains and also altering their beliefs about practice. Firstly, they became more aware of the need to teach thinking explicitly with a focus on the process of thinking and learning. This contrasted with earlier beliefs of the teacher as an imparter of knowledge. Thus their view of their role in this process shifted from that of information giver to that of facilitator and encourager. Staff tended to be less preoccupied with written products of learning and instead were more focused on the process of thinking in curriculum contexts. Munro (1999) displayed a corresponding effect with teacher's knowledge and beliefs about the teaching of thinking evolving. Blagg (1991), likewise, indicated a development of more positive attitudes of staff in this area. Finally, staff became more self-reflective and evaluative of their professional practice. This may be linked to the time invested in the ACTS intervention training. Alternatively it could mirror the metacogntion actively encouraged in ACTS classrooms. It is also indicative of the connection between changing teacher beliefs on alternative classroom practice (Calderhead, 1996). These findings confirm Hypothesis 5.

Future Considerations

A key limitation, identified by groups, was that of time needed to plan and implement intervention lessons. This included the "*extra time needed to plan and evaluate lessons at first*" (Amir, Question 2, Theme 1, Code 1.4), alongside the "*time to teach the thinking within lesson/timetable constraints*" (Helen, Question 2, Theme 1, Code 1.4). Despite this additional time commitment teachers noted that although "*length of lessons increased – children's learning now correspondingly greater*" (Mary, Question 2, Theme 1, Code 1.4). This reiterates findings made McGuinness et al. (1997) in the first exploratory study into ACTS. These issues are addressed in later section in relation to EP practice.

An additional issue raised by staff was the value in collaborating and sharing practice from ACTS in a group training setting. Questionnaire responses indicated that practitioners valued the opportunity to share resources and plan collectively. In particular staff believed long term planning and inclusion in the School Development Plan would be crucial for the success and implementation of any thinking skills intervention. This would mean any framework was embedded across key stages and curriculum areas to support the transfer and maintenance of thinking skills over time. The possibilities of this are reviewed in the implications for EP practice section.

Implications for the Knowledge Base in Psychology

The study contributes to the psychological knowledge base along a number of key dimensions in a distinct manner. This section aims to outline such areas in terms of both theoretical developments and clinical implications in the thinking skills sphere.

Theoretical Implications

This project adds to the theoretical debate in thinking skills research. Firstly it confirms that children's cognitions can be developed following a thinking skills intervention of a two year duration. In this research qualitative data from students and staff alike evidenced changes in children's use of thinking vocabulary and explicit use of thinking skills. Such findings are evident by others in the field (Adey et al., 2002; McKinstery & Topping, 2003; Trickey, 2004). This information can be used in the debate surrounding learning and development (Adey & Shayer, 1994; Anderson, 1992; Case, 1978).

Learning is regarded as a consequence of action and is independent of maturation whilst development is considered a slow irreversible process and related to maturation (Adey et al., 2002). Many in the field espouse thinking skills interventions promote cognitive change linked to development as opposed to mere learning (Adey et al., 2002). For such a position to hold true, evidence of transfer, retention and longevity would be required (Blagg, 1991; Sternberg & Bhana, 1986). This study did indicate transfer of effects, from within the classroom to the social settings of the playground and out of school, from the qualitative responses. Moreover such data demonstrated an

impact on behaviour and self perceptions. Longevity and retention of effects, however, requires more long-term follow-up to investigate stability over time. Such work contributes to the field in supporting suppositions that thinking skills can potentially be developed (Adey et al., 2002; Lipman et al., 1980; McKinstery & Topping, 2003; Topping & Bryce, 2004; Trickey et al., 2004).

The research confirms findings in the existing empirical evidence base pertaining to generalisation of effects. To date research has suggested cognitive changes can be made for both secondary (Shayer & Adey, 1993; McKinstery & Topping, 2003) and infant aged pupils (Yr 1) (Adey et al., 2002). This study makes a unique contribution in examining the effects in the primary age range. The piece confirms evidence that no differential impact is noted from interventions on a range of social variables such as ethnicity, electoral ward standing of the schools and free school meals (Adey et al., 2002; Blagg, 1991; McKinstery & Topping, 2003). This supports the view that such cognitive interventions are applicable to all pupil populations (DfEE, 1999; Higgins et al., 2004).

The project informs the discussion surrounding the theoretical bases of interventions currently available. Firstly it provides a distinct contribution in evidencing that an infusion intervention can make a positive difference on student and staff development. This is the only confirmation to date using both quantitative and qualitative data. Such information suggests infusion could be a potential intervention alongside those of a subject specific or bolt on approach.

With regards the theoretical underpinnings of varying approaches, this study confirms many of the pertinent principles. The ACTS intervention actively uses the notion of metacognition (Baird, 1984; Brown, 1987) throughout the sessions in line with programmes of cognitive acceleration (Adey et al., 2002). Furthermore principles of social construction are used in the actively thinking together element of the lessons. Such ideals link to Vygotsky's (1978) concepts whereby knowledge develops as part of a social process. This has been pinpointed as crucial by many in the area (Higgins et al., 2004; Joiner et al., 2000; Light & Littleton, 1999; Mercer, 1995; Wegerif et al., 1999). In such a manner the study emphasises the relevance of these cognitive theories. It does not, however, establish the relative importance of each. This would require a sophisticated study of independent comparison as discussed in the later section on future research.

The relationship between cognitive development and academic achievement is validated by this research in line with earlier work (Adey & Shayer, 1993; Lipman, 1991). Prior work reported that school attainment could be enhanced through the promotion of more efficient processing skills, for example the use of cognitive acceleration (Shayer & Adey, 1993; Shayer & Beasley, 1987). Such findings emphasises the need for time before gains became apparent (Shayer & Beasley, 1987). This study supports such suppositions as attainment was reported to improve by staff although alterations were not evidenced in quantitative measurements.

Theoretical links between attainment and self-belief of learners (Tuckman, 1999) are supported in this study. Such a relationship was evidenced by pupil scores on the Myself-As–A-Learner Scale (Burden, 1998) compared to those on the Cognitive Abilities Tests (Lohman et al., 1993). It appeared that pupil perceptions about their learning and abilities were connected to their attainment on the CATs. Such a fact has been illustrated elsewhere (Thorndike & Hagen, 1986). This supports theoretical perspectives pertaining to children's self concepts and its connections to performance. Furthermore the piece adds to the study of students self-perceptions in relation to tracking changes linked to an intervention. It also examined pupils' views of themselves as learners and compared this to teacher perceptions as monitored by the Taxonomy of Problematic Social Situations (TOPS) (Dodge et al., 1985). This extends work examining the impact of teachers on pupil self-perceptions of themselves as learners (Burnett, 1999).

Finally the research added to existing research by examining pupil perceptions to learning in the primary age range. Such a domain has been hereto untouched: "*few studies have investigated these phenomena using elementary school students and psychometrically sound instruments*" (Burnett & Proctor, 2002) (p. 325). This study sought to examine pupil perceptions in relation to self-concept using relevant and rigorous tools contributing to research in this realm (Burnett & Proctor, 2002). This has been highlighted as an area in need of scientific scrutiny for some time (De Corte, 2002; McGuinness, 1999; Watkins et al., 2001; Wilson, 2000).

Clinical Intervention Implications

The current study contributed to thinking skills research given its increased rigour in research design. In this manner it informs the area of intervention evaluation studies. Firstly the piece used a *control group* beyond that of the *no treatment control* prevalent in prior work (Adey & Shayer, 1993; Blagg, 1991; DfEE, 1999; Lipman et al., 1980).The use of a *waiting list control* group which was matched to the *experimental group* ensured expectancy could be controlled for. Moreover, such group allocation considered matching schools along demographic characteristics which had been unaccounted for in previous projects. Earlier studies had investigated interventions with staff who self-selected to become part of the project including ACTS, CASE and IE. This study removed such a factor and by selecting schools eradicated the role of staff self-selection. Finally earlier work had failed to detail the rationale and relevance of measures used (Gorodetsky et al., 2000). This study addressed such criticisms by justifying and detailing measurements employed. In these various ways it extended rigour into the routines of thinking skills research and readdressed earlier limitations through the scientific collection of empirical evidence (Higgins, 2002; Stoiber & Kractochwill, 2000).

The piece of research provided a unique contribution to cognitive intervention studies by being the first to examine an infusion intervention using qualitative and quantitative methodologies (Higgins et al., 2004). To date the only evaluation of an infusion approach was in Northern Ireland with teacher subjective feedback alone (McGuinness et al., 1997). This study distinctively

adds to this thorough the use of multiple measurements and methodologies. In a similar vein the piece is original in evaluating the intervention in everyday classrooms, an omission of earlier work (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000). Similarly it makes the focus a primary setting hereto ignored by the majority of research (McKinstery & Topping, 2003). Furthermore measures of both student and staff development across a number of domains addresses issues of transfer untouched in earlier work (Blagg, 1991; DfEE, 1999; McGuinness, 2002; Wilson, 2000).

The study also confirmed areas worthy of further scrutiny and development in the thinking skills domain. Issues of implementation and intervention integrity were identified in the piece which has implications for research and practice alike. Future studies will need to address such matters with monitoring in place to ensure the espoused intervention is that in use with frequency across the curriculum determined. Organizational issues relating to time, to both plan and implement interventions were, likewise, indicated as critical in line with similar studies (McKinstery & Topping, 2003). These wider systemic issues were also evidenced in relation to the impact of school and teacher on the success of an intervention in context. Other work has yielded contradictory information as to the relevance of both variables and as such offers an important area of further investigation (Adey et al., 2002; Blagg, 1991; McKinstery & Topping, 2003).

Finally aspects of advanced analysis were noted by the current research. The positive impact on bilingual learners, for example, was identified by many teachers. This confirms the opinion that cross-cultural comparisons in relation to such interventions are warranted (Dart et al., 2000a). Correspondingly the effect on Special Educational Needs was regarded as powerful with further examination needed (McKinstery & Topping, 2003). Specific issues such as maintenance, transfer and longevity of effects still exist and work is required given there significance in this sphere. Lastly given the multifaceted nature of thinking and the limitations in measuring it, as demonstrated in this study, additional work is justified in this area (Kember et al., 2004).

Implications for Future EP Practice

Educational Psychologists are well placed on the interface of research and practice to translate knowledge regarding children's cognitive development into the educational environment. This section delineates the potential part EPs can play in this evolving area. In particular this involves the application of psychological knowledge to professional practice at a student, staff, school and EPS level.

Student Level Applications

Educational Psychologists have an understanding of children's development and special educational needs (Baxter, 2002) which can be used within the specialist sphere of thinking skills. In light of this research, EPs could be instrumental in transferring such conceptual knowledge into everyday usage.

In particular piloting the measures, materials and approaches to cognitive instruction highlighted to ascertain application. One focus could be to use and critically evaluate the measurement devices designed to tap into thinking skills and metacognition. This is pinpointed in research as a distinct area of need as McKinstery and Topping (2003) argue "*exploration of adequate and practical tests of thinking is required*" (p.214). Indeed this research indicated the need for more measurement devices. EPs could use their specialist knowledge of assessment techniques to inform the design and development of such tools. This is particularly pertinent given the need to obtain the pupil perspective as outlined in Every Child Matters (H. M, Treasury, 2003).

Practitioners could critically evaluate current measurement devices from a scientific perspective with a range of students. Comparisons between measures could be explored from technical detail to usage. Self-concept measures from Burnett (1994; 1991) could be compared with Approaches to Learning scales such as Bigg's "Learning Process Questionnaire" (LPQ) (1987). Moreover computer assisted technology such as the "Pupil Attitudes to Self and School" (PASS, Williams et al., 2003) could be investigated to gauge if this form of administration is preferable to students. EPs are ideally located to critique measures through use in the applied setting to determine appropriateness with different ages, special educational needs or cultural populations (Dart et al., 2000a). Key issues could be evaluated such as a measure's validity in distinguishing between surface and deep level learning (Burnett & Proctor, 2002; Kember et al., 2004) or the impact of age (Nisbet, 1993; McKinstery & Topping, 2003). Such work could then inform the

development of more finely tuned methods for specific pupil populations for as Costa (2001) highlights "*we must seek additional means of determining growth in intellectual abilities*" (p.326). Outdated measures could also be revised and adapted to today's education system (Kember et al., 2004).

Educational Psychologists could be influential in examining unanswered practice questions highlighted by this research in working with students. Action research with colleagues to explore existing interventions in schools and in particular the impact on pupils could be examined (Fox & Sigston, 1992). As Baxter (2002) indicates "*a potential role for EPs is that of contributing to the initial design, project monitoring and evaluation of pilot studies provided locally*" (p.68). In this respect the effectiveness of such projects in practice could be ascertained, especially with reference to areas of ambiguity highlighted in this and other research linked to learners (Higgins et al., 2004; McGuinness, 1999; Topping, 2001; Watkins et al., 2001).

The generalisation of higher order cognitive thinking skills across time and context could also be examined by EPs (Blagg, 1991; McGuinness, 1999; Nisbet, 1993). The concepts of "*high road* " and "*low road*" transfer could be explored (Salomon & Perkins, 1989) alongside the durability of effects (Adey et al., 2002; McKinstery & Topping, 2003). The issue of optimum age of delivery could be investigated (Nisbet, 1993). Such areas could be linked to models of children's cognitive development to ascertain appropriate interventions according to different stages (Inhelder & Piaget, 1958). Moreover comparison of different types of thinking skill intervention with

different ages of students could be examined (Higgins et al., 2004). For example would Paired Thinking be more appropriate for primary or secondary aged children (McKinstery & Topping, 2003) and are infusion methodologies more appropriate for primary settings (McGuinness, 2003).

Educational Psychologists could evaluate the impact of thinking skills interventions on all aspects of the learner from the emotional to the academic (De Corte, 2002; DfEE, 1999; McGuiness, 2002; Watkins et al., 2001; Wilson, 2000). Such work could build on this research in examining multiple learner outcomes. The impact of interventions on self-efficacy could be explored alongside self-esteem (Ames, 1986). Issues of self-concept as a learner could be thoroughly tracked over time to gauge stability and development in relation to interventions (Burnett & Proctor, 2002). Likewise the impact of thinking skills instruction could be determined across the age range in relation to changing self-concepts (Burnett, 1999; Dart et al., 2000b). Comparative effects of interventions from infusion to enrichment could also be examined to determine optimum delivery according to student's development (Coles, 1993; DfEE, 1999; Higgins et al., 2004). In particular whether as maturity develops students respond more favourably to peer mediated (McKinstery & Topping, 2003) or subject specific approaches (Adey & Shayer, 1993).

Staff and School Level Applications

From a practice perspective EPs can use specialist knowledge from thinking skills research to inform staff and schools development alike using both expertise and process consultancy (Fox & Sigston, 1992). EPs have the understanding of thinking skills instruction and can use this to support schools in implementing thinking skills interventions. In terms of the objectives of schools as organisations, EPs can provide the scaffolds required to develop the curriculum to encompass thinking skills strategies. Such pedagogical change in terms of both policy and practice will be necessary (DfEE, 1999; Higgins et al., 2004). EPs could systematically work with schools to grapple with issues of curricular design, choice of materials and pedagogy (McGuinness, 1999; 2003). They could provide information as to the placement of programmes within the existing curriculum alongside how best to support transfer and maintenance, key issues raised by schools in this study. As McKinstery and Topping (2003) point out "*Educational psychologists are well placed to consult regarding the effective organisation of such initiatives*" (p.213). In this manner EPs can inform organisational development and school improvement in relation to thinking skills interventions. Furthermore they can assist schools in selecting approaches in the burgeoning market of commercially constructed thinking skills packages (TES, 2002).

At a staff level EPs can use INSET and process consultancy to affect change in classroom practice in connection to thinking skills interventions. As this study and others have demonstrated staff development is critical to the

implementation integrity of a thinking skills intervention (Adey et al., 2002; Burnett & Proctor, 2002; McKinstery & Topping, 2003). As Adey et al (2002) argue "*such pedagogical skill cannot be delivered by printed materials alone, but requires a carefully designed professional development programme*" (p.12). EPs could be influential in using adult models of learning to support staff in taking on new techniques of teaching (Binsted, 1982; Knowles, 1984; Kolb et al., 1984). Such a programme would need to encompass both theoretical knowledge and experiential translation of such concepts, both of which an EP can provide. In particular the process of learning (Biggs & Moore, 1993) and the role of the learning environment (Burden & Fraser, 1993; Burnett & Proctor, 2002; Fraser et al., 1989) could be considered. This in turn would support intervention implementation for as Adey et al (2002) indicate "*effective delivery of these activities depends on the teachers having a good understanding of the underlying theory and much practice*" (p.12). As this study evidenced staff valued the training and review days to receive, share and evolve techniques in a supportive manner. EPs could provide the ongoing training and staff development for schools in maintaining and monitoring thinking skills instruction in context.

EPs can also work with schools to determine the variety of processes best suited to enhance children's thinking and learning. This can draw on the empirical evidence base of which types of interventions generate most favourable results for specific students. This could look beyond the mainstream school context, an area omitted in existing literature (McKee & Witt, 1990; Ysseldyke & Christenson, 1987). Moreover advancements in

technology could be examined as the European Agency for Development in Special Educational Needs (2001) highlights *"what needs to be developed are methods on how to use ICT as a pedagogical aid in the teaching of all pupils"* (p.213). In this manner ICT could be used not only to measure change but support cognitive development in relation to thinking. In particular resources such as Thinker Tools (White & Frederiksen, 1998) and interactive video discs (Vye et al., 1997) could be applied in context. Such devices embed thinking in computer mediated learning to provide not only real life application but also systematic and explicit reflection on thought processes. Moreover a range of researchers evidence the ideal opportunity that computers provide, in supporting children's collaborative learning and associated cognitive development (Crooks, 1994, 1995, 1996; Dillenbourg, 1999; Littleton & Light, 1999; Wegerif & Scrimshaw, 1997). Finally EPs can support schools in the evaluation of intervention effectiveness across a range of variables in small scale action research.

EPS and LEA Level Applications

At an LEA level EPs could be instrumental in policy and practice development linked to thinking skills evolution throughout education. In particular EPs could assist in the interpretation and implementation of the government's 5 year strategy for children and learners (DfES, 2004). This would advance and accelerate the pace of reform in teaching and learning. At a primary level EPs could scaffold school networks in collaboratively working together to embed thinking skills within the curriculum. Whilst at the secondary and college level EPs could support organisational development

to ensure all learners were enskilled for the 21st century work place. This would involve EPs working systemically in the LEA to ensure government policy was enshrined and translated meaningfully into educational practice.

At the level of local education authorities psychologists can be instrumental in both implementing and evaluating thinking skills programmes. As work in Clackmannanshire indicates, the psychologist's role was to set up and evaluate a county wide initiative based on Lipman et al's 1980 Philosophy for Children (Trickey, 2004). This included providing staff INSET, critiquing the research literature and evaluation using video analysis, standardised tests and questionnaires (Trickey, 2004). Such work was justified in providing the greatest contribution to most children in line with MacKay's 1999 self evaluation. Indeed as Higgins et al (2004) indicate "*research could be commissioned to establish what is both effective and efficient (particularly what is cost effective) in terms of thinking skills interventions* "(p.46). It is at such a level that EPs can potentially have the widest impact. Such research could occur in conjunction with the Learning and Skills Council (LSC). Outcomes and evaluations could be connected to the 5 Outcomes of Every Child Matters (DfES, 2003) for LEA accountability. In this manner EPs would be ensuring the rigorous use of research and evidence based practice to explore the best way forward in this domain (Baxter & Frederickson, 2005). Furthermore it would support learning and thinking of all young people form 0-19 in keeping with the governmental push for key skills development (DfES, 2004).

Recent legislation necessitates substantial shifts in the practice of Educational Psychology at an LEA level with the push for Integrated Children's Services (DfES, 2003). Given the restructuring and cultural change such moves may warrant it may be that the application of thinking skills will be useful at an EPS level. Indeed as Jensen et al (2002) demonstrate, altering patterns of thinking at both the level of individual practitioner and services as a whole may promote such change. As a profession the expertise and knowledge we possess regarding thinking could be used proactively to carve out our future roles accordingly. This could use strategic tools such as force field analysis or the component technologies of learning organisations (Senge, 1993). In this way the psychological knowledge of thinking and change could be utilised within the discipline to constructively evolve into the future (Jensen et al., 2002). As Baxter (2002) points out EPs "*are well placed to take on confident, new, strategic roles*" (p.65) and the application of creative thinking to ourselves may be one such vehicle for our continued development.

Implications for Further Research

As McKinstery and Topping (2003) pinpoint "*Thinking skills currently have high profile*" (p.199). Indeed this is reflected in the plethora of packages aimed to enhance thinking commercially available. Despite claims to make thinking more effective this is not borne out by research, with evaluations remaining inconclusive (DfEE, 1999; Gorodetsky et al., 2002; Wilson, 2000). There is considerable debate as to the impact of varying approaches with a need for more empirical evidence (McGuinness, 1999; Watkins et al., 2001).

This study, likewise, indicated a number of limitations and unanswered practice questions remaining. Educational Psychologists are uniquely placed to engage in theory driven research to make connections between practice and research transparent (Bond, 2002; DeCorte, 2002). This section illustrates the role of EPs in engaging in more scientifically rigorous research to extend the current evidence base by addressing practice questions and pursuing new avenues for development.

Extension of Current Research to Redress Limitations

As the critique of this study and others demonstrate thinking skills studies have tended to use inadequate control groups in comparison with the intervention, an issue which has dogged the research for some time (Sternberg & Bhana, 1986). In order to detail an effect linked to a specific intervention a wider number of control groups would be required (Greenhalgh, 2001). Such a design would ensure groups were comparable in all important aspects apart from the intervention under study. This would require large scale work with an *experimental, waiting list control* and *no treatment control* group from within the same school. Such a method would ensure the impact of a range of variables, including the school and student characteristics, were controlled for. Furthermore randomisation of group allocation within the school would need to be masked to researchers and not influenced by staff or school selection. Such groups would then need comparing across a range of schools to discredit any confounds. In addition schools and students would be randomly selected from a range of geographical locations to form a representative sample of the national

population. This would also resolve earlier limitations levelled at research pertaining to the fact that interventions occur in optimal learning environments (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000). Such a sample would focus in on the impact of the intervention in ordinary classrooms with randomly selected staff and students.

The size and scope of studies would also need extending in terms of time frames for as McKinstery and Topping (2003) argue "*more research is required for longitudinal studies before any firm conclusions can be drawn*" (p. 214). In this piece, for example, follow up at the end of Year 6 and then into Year 7 would have been an insightful method to gauge the transfer of effects to the Secondary school. Moreover long term follow up of students over significant periods of time could inform the debate regarding both optimum age and issues of transfer (Adey et al., 2002; Blagg, 1991; Nisbet, 1993; Wilson, 2000). In particular the periods of transition could be examined alongside the arguments for the benefits of early intervention (Adey et al., 2002).

Future research would require more scientific rigour in terms of both the intervention implementation and selection of measurement devices. With regards to thinking skill interventions applied, copious detail would be needed to allow for a replication, a limitation levelled at both this study and others in the field (Adey & Shayer, 1993; Blagg, 1991; De Bono, 1992; McGuinness et al., 1997). Likewise monitoring of the integrity of the implementation would be critical to ensure a true use of the intervention in context (McKinstery &

Topping, 2003). In addition this would yield relevant information as to the frequency and placement of thinking skills interventions within the curriculum. In terms of measurement devices, new work would need to use multiple devices to ascertain outcomes. A review of the research and this project indicates a restricted use of tools to tap into learner variables successfully. As Burnett and Proctor (2002) detail "research *in this area has been hindered by a lack of suitable instruments for gathering data*" (p.325).

In future studies, greater attention needs to be focused on the triangulation of measures to evidence change on a range of outcomes. These will need explicit reliability and validity statistics alongside measurement over time from various perspectives from self evaluation to teacher and parents (Costa, 2001). Moreover, the combination of quantitative and qualitative design will illuminate the mechanisms of change. For example in this study triangulating subjective participant feedback with questionnaires, focus groups and semi-structured interviews would be a distinct improvement.

Addressing Unanswered Practice Questions

Throughout thinking skills research there have been numerous unanswered practice questions in need of extended analysis (DfEE, 1999; Higgins et al., 2004). One of particular pertinence pertains to identifying the critical factor which promotes change in children's thinking (Adey et al., 2002; McGuinness et al., 1997). Indeed in this study the role of language in developing thinking skills was evidenced in teacher report. Also the impact of thinking co-operatively in groups was highlighted by students and staff alike. Future

studies will need to isolate the active ingredients in thinking skill interventions which act as the mechanism of change. Such developments link clearly to the theoretical underpinnings of programmes. Work could examine, for example, the role of metacognition (Baird, 1984; Blank, 2000; Brown, 1987; Doherr et al., 2005) in making thinking about thinking an explicit process to support the enhancement of more general thinking processes (De Corte, 1990; McGuinness, 1990; Saloman & Perkins, 1989). This could be compared against Vygotskan principles of social construction, in that children's thinking is developed through a process of interaction in the social environment (Vygotsky, 1978). The importance of peer collaboration on cognitive development has been supported by others (Adey et al., 2002; Joiner et al., 2000; Light & Littleton, 1999; Mercer, 1995). Research could attempt to determine the impact of such factors on children's thinking skill development to determine the crucial component through alternative intervention designs. Such work would illustrate the contribution of various elements of effectiveness.

Investigation surrounding positive effects could, likewise, be examined for various pupil populations. This could range from disadvantaged social settings (Adey et al., 2002) to special educational needs groups (McKinstery & Topping, 2003) and special school populations (McKee & Witt, 1990; Ysseldyke & Christenson, 1993). Moreover cross-cultural and age comparisons would be needed (Dart et al., 2000a).

An area fraught with difficulties in the domain of thinking skills is the actual process of gauging changes in children's thought processes. As McKinstery and Topping (2003) explain "*measuring improvements in thinking is difficult*" (p.205). As this study indicated the paucity of assessment tools in the field limits the extent to which alterations in children's thinking can be ascertained. Such an area has been at the heart of difficulties for some time (Blagg, 1991). An essential domain for new research needs to be the development of more rigorous devices to tap into children's thinking processes (Bond, 2002; De Corte, 1990). Such work would need to ensure a valid and reliable means to measure cognitions across the chronological age range to parrell maturity and development. Substantial work exists in the context of higher education and needs extending down to school aged children (Burnett & Proctor, 2002). Specific devices could be designed from activities using Piagetian protocols of stage of development (Inhelder & Piaget, 1958) to questionnaires measuring metacognition (Goos, 1999).

The possibilities for practice also need addressing in triangulating data from such measures with alternative perspectives from teachers and parents systematically over time (Costa, 2001). Links to the intervention and the environment could also be made using tools to distinguish between deep and surface approaches to learning(Dart et al., 2000b). This could explore the model of teaching utilised and impact of individual approaches to thinking and learning. Finally in the construction of such measurement tools we may need to rethink accepted and employed testing procedures. As Kember et al (2004) argue "*many psychological constructs are complex in character so*

would be more widely portrayed as multidimensional" (p.272/3). This calls into question traditional procedures for instrument development which uses reliability alone. Such developments may be necessary given the multi-faceted nature of thinking and the need to capture these constructs comprehensively.

Future Research Directions for Development

A range of research possibilities exist stemming from the omissions in this study and the literature to date. Firstly, there is scope for further focused work in ascertaining the role of emotions and dispositions of children to thinking (De Corte, 2002; McGuinness & Nisbet, 1991; McGuinness, 1999; Watkins et al., 2001; Wilson, 2000).This study used the Myself-As–A-Learner scale (Burden, 1998) alone and could have complimented this with other measures such as the Burnett Self Scale (BSS, Burnett, 1998). Such investigations could concentrate on children's self-concepts further and how these could be enhanced through thinking skills interventions to offset decline in the elementary years (Burnett, 1996). Correspondingly *"links between generalised self-concepts and educational outcomes"* could be examined (Burnett & Proctor, 2002) (p.232). Relationships between various aspects of self-concept could also be explored in relation to both the thinking skill intervention type and comorbid changes in children's approach to learning. Gender and cultural differences, likewise, deserve further exploration in this sphere (Dart et al., 2000a).

An additional avenue for development is the realm of the social situation and significant others in supporting cognitive development (Crooks, 1995; Joiner et al., 2000; Light & Littleton, 1999; Mercer, 1995). As McKinstery and Topping (2003) discuss "*much of the research literature is concerned with teacher directed instruction in thinking skills, and much of it relates to higher education rather than schools*" (p.200). Future work could redress this issue by examining the role of others to support children's development of thinking. Firstly alternative interventions to the principally teacher led approaches could be explored with reference to peer mediators or paired tutoring of thinking skills (McKinstery & Topping, 2003; Topping & Bryce, 2004). Emerging work in this domain is indicative of positive affects for both tutors and tutees (op.cit). Moreover the potential part to be played by parents and carers could be investigated as "*there is practically no research literature on parent or volunteer tutoring of thinking skills*" (McKinstery & Topping, 2003) (p.200). Studies into this would also inform the assessment of changes to thinking as noted from a variety of sources systematically over time (Costa, 2001). The role of computers in terms of providing a focus for children's collaborative learning could also be examined (Crooks, 1998; Dillenbourg, 1999; Littleton & Light, 1999; Wegerif & Scrimshaw, 1997).

Finally the wider context could be considered to ascertain the impact in a whole school in a given thinking skills intervention (McGuinness, 2002). As this study demonstrated, through head teacher questionnaires, the ACTS intervention made a distinct difference to classes within a school. Additional research to gauge organisational change in schools in relation to thinking skill

interventions would be an area of application meriting work. Such studies could assist in determining the key practice elements required for the successful implementation of a thinking skills intervention in the school setting (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000). This could draw on work including that of powerful learning environments (De Corte, 1980) and the need for professional development programmes to support staff (Adey et al., 2002).

Staff skills could be specifically studied to ascertain if teacher effectiveness is enhanced as evidenced in earlier work (Blagg, 1991; Munro, 1999; McGuinness et al., 1997; McKinstery & Topping, 2003; Stewart & Smardon, 2002). The use of objectively selected samples and measurement devices would enhance work in this area. It could ascertain the features for support and change in terms of teacher knowledge and skills over time and contexts to highlight success factors. Longitudinal work could also compare the various types of thinking interventions in similar schools to ascertain the most effective (Coles, 1993). As Higgins et al (2004) argue "*there is a clear need for more comparative studies between different types of intervention, and between thinking skills approaches*" (p.6). Research could ensure that interventions are based on an empirically rigorous evidence base, which is much needed within the education system (Higgins, 2002; Higgins et al., 2004; Gorodetsky et al., 2002; Stoiber & Kratochwill, 2000).

Chapter 5: Evaluation of Study

This study offered the first long term evaluation of an infusion intervention in the educational environment. Its credibility is reduced, however, by a number of methodological and measurement flaws which will now be outlined. The following section aims to critique the study using critical evaluation checklists developed by Greenhalgh (2001); Rudestam and Newton (1992) and Connor (1997). Initially methodological issues will be discussed followed by matters pertaining to interpretation and analysis.

Methodological Considerations

The study had a detailed theoretical base which drew on a range of psychological theories and was established on existing research in the field (McGuinness et al., 1997; McGuinness, 2002). It was original in adding to the empirical evidence base in being the first long term evaluation of an infusion intervention using both qualitative and quantitative methods (Higgins et al., 2004). Criticisms exist, however, in terms of the sample, design and measures utilised.

Sample

With regards to the sample the study attempted to examine a thinking skills intervention in natural circumstances with both students and staff alike. This was original from inception and addressed earlier criticisms levelled at research conducted only in optimal learning environments (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000). Moreover, earlier limitations of

self-selected staff participants were avoided by selecting schools instead. In terms of the sample of schools the original cohort represented a cross section of contexts from Midshire. Evidence of recruitment bias, however, existed as only schools who wanted to participate in the study volunteered. In this manner the sample of schools could be skewed as only those who were potentially motivated in developing thinking skills agreed to take part. Similarly in the selection of schools, SAT scores were considered to ensure a range of different schools. As Adey et al., (2002) indicate a more controlled and scientific sampling device would be required in future as such a measure may be biast.

An additional limitation was the incomplete randomisation of schools to the *experimental* or *waiting list control*. In attempting to match schools between the two groups, experimenters were influential in allocation. This may have introduced systematic differences linked to selection bias into proceedings (Sackett et al., 1991). In future researchers would need to be masked to this process. The adequate matching of schools between groups was also restricted following the choice of some schools to remove themselves from the study. This meant demographic characteristics between the *experimental* and *control* groups became skewed reducing meaningful comparisons which could be made. As the initial descriptive statistics demonstrated, significant differences between the *control* and *experimental* group were found on the baseline CAT scores. In the original research design 12 schools were selected, 6 of which would form the *experimental* group and 6 for the *waiting list control*. Once the study was underway, however, 2 schools felt unable to

commit to the demands of the project. This meant that the matching of schools between *experimental* and *control* was compromised in terms of both numbers and composition. In the future these groups would need closer matching to avoid significant differences at baseline. Moreover, it became apparent during the project that some schools were implementing other thinking skills interventions. Such a fact contaminates findings and meant these schools were removed from the data set.

Finally the study made no attempt to detail how representative the student and staff sample involved were of the national population. In future finer detail regarding the sample would be needed to eradicate any bias and inform the level of extrapolation. Clear inclusion and exclusion statistics with accompanying rationale would, likewise, be informative in making generalisations. In particular more specific detail on teachers would be important given the potential impact evidenced in other studies (Adey et al., 2002).

Research Design

The study had clearly stated research questions which connected to the hypotheses and the empirical evidence base examined. Several issues occur, however, in relation to the adequate use of an intervention comparison. Firstly, the study failed to provide a sufficient number of control groups. The use of an *experimental group* and a *waiting list control* allowed for some comparisons to be made and expectancy controlled for. A *no treatment control* would be necessary in the future, to control for age and

maturation alike. In this manner inadequate control in the study meant a range of extraneous variables were unaccounted for. Such matters were compounded by insufficient randomisation of groups and limited detail regarding participants. Indeed the study failed to examine baseline differences between the groups explicitly. In future key features of both groups would need ascertaining so as to allow the reader to make direct comparisons accordingly. This would indicate to what extent the groups differed in all important variables alongside the intervention under study.

The research did not provide sufficient information pertaining to the intervention which, meant replication was impossible. Furthermore there was no monitoring as to the ACTS intervention implementation other than that of teacher self report. In this way the study did not state the frequency nor placement of the thinking skills instruction across the curriculum. This calls into question the integrity of the intervention across both schools and staff. Consequently assumptions are made, in that what teachers espoused to be doing reflects what they are actually doing. In future monitoring the integrity of the intervention could be achieved through visits and checklists (McKinstery & Topping, 2003). Similarly data on the frequency of lessons could be recorded alongside observations and videotaping (McGuinness, 2003; Trickey, 2004). This would provide a check on how teachers were both interpreting and implementing the intervention accordingly. Likewise it would offer information as to the optimum levels and location of thinking skills lessons across the curriculum.

Outcome Evaluation- Measures Utilised

One of the study's strengths was the fact that it was one of the first attempts to measure multiple outcomes in terms of the pupil participants. Issues of validity and reliability of measurement devices were discussed and selection justified unlike earlier studies (Gorodetsky et al., 2002). Despite this the piece failed to operationalise variables sufficiently with limited triangulation of evidence. Such limitations pertain to both qualitative and quantitative measurements alike.

In terms of quantitative measures the majority of devices, although, objectively validated, provided a crude insight into children's skills. The use of the Cognitive Abilities Tests (CATs) to tap into pupils' thinking skills is one such example. As Costa (2001) argues "competency *may be demonstrated in a single test, intellectual effectiveness is demonstrated by sustained performance in a variety of situations"* (p.326). The use of the CATs was unlikely to reflect changes in children's thinking skills and is limited by performance at a snapshot in time (Coffman, 1980; Lazar-Morrison, 1980). In future more specific examination of thinking skills strategies could be used such as the activities in Paired Thinking (McKinstery & Topping, 2003). Likewise, data collected systematically over time through observations from teacher and researcher would not only be more representative and realistic but illustrate transfer of thinking skills. Materials tapping into the pupil and parental perspectives could also supplement such evidence (Costa, 2001).

Other measures utilized in the study were similarly flawed. The use of the Myself-As–A–Learner scale (Burden, 1998) was limited in providing a general perspective of learning in school. A more sophisticated tool to use could have been the PASS (Pupil Attitudes to Self and School, Wiliams, et al., 2003). This would provide more informative and detailed data on both self-esteem and self-perceptions as a learner. In addition the use of the Taxonomy of Problematic Social Situations (TOPS) (Dodge et al., 1985) used only the teacher perspective as an indicator of behaviour. Improvements would be needed in terms of triangulating this measure and others with additional evidence. For example the TOPS could be substantiated with observations, self-evaluations and parental report which could reduce subjectivity. Moreover measures of attainment from a variety of sources over time could evidence generalisation and maintenance of skills across contexts (Costa, 2001). The combination of self evaluation, observation from teacher and researcher alongside parental report would triangulate evidence across a number of variables (Costa, 2001).

Procedurally, the administration of assessments at the beginning of the academic year may have influenced results. It may be that pupils' performance was an underestimate of abilities given children had only just returned from the summer break. Furthermore the fact that experimenters conducted the assessments may have impacted on pupil outcomes. The majority of assessments occurred on one school day, which may mean the effects of fatigue and lack of concentration confound results obtained.

Qualitative measurement devices were also flawed which undermines the study's credibility. Firstly the use of thematic analysis, although, rationalised did not provide the scientific rigour necessary for a study of this scope. Furthermore the use of questionnaires for pupils, teachers and head teachers provided only one method of measurement which did not control for the expectancy of improvement or the Hawthorne effect. Such techniques may have limited the richness of responses obtained and could have been avoided through the additional use of semi-structured interviews and focus groups. Issues of questionnaire construction and administration with respect to each participant group were also problematic and are now reviewed.

The advantage of using teacher questionnaires from Professor Carol McGuinness meant this research paralleled that of a sister project in Northern Ireland. However, there were issues with validity and reliability that were not adequately addressed. The actual time at which questionnaires were completed may have also influenced staff responses. Firstly experimenters were present which might have positively skewed data collected. The fact that the questionnaire was completed on the final review day of training may also mean individuals were influenced by the majority view voiced predominantly on that day. The analysis and interpretation of results may have been coloured by experimenters' subjectivity particularly given their involvement in the training, implementation and evaluation of the ACTS intervention. This was, however, controlled for by outlining the steps of data collection and analysis (Yin, 1994) alongside making raw data

available so findings and interpretations could be examined (Erlandson et al., 1993).

The head teacher questionnaire was also a self report measure completed at the end of each academic year. It was constructed by the experimenters to ascertain the impact of ACTS at a school level and may therefore have been skewed in its scope. It suffers from the same reliability and validity restrictions as the teacher questionnaire as it represents a non-standardised measure. Despite this it provides the head teacher perspective on the impact of ACTS, an area untouched in prior work (DfEE, 1999; McGuinness, 2003).

In terms of pupil participants the experimenters involved in the project constructed a measure used to elicit children's views. This limits the objectivity of the device as it may be influenced by researcher's agendas. The validity and reliability of the measure were not ascertained on a standardised sample prior to use. Likewise a pilot study did not occur to remove any misleading or ambiguous questions. Despite these difficulties the measure did attempt to identify and include the child's perspective, which has hereto been ignored in research to date (DfEE, 1999). The quantitative aspect of this questionnaire was disregarded in analysis as the questions were considered misleading as they contradicted the ethos of the ACTS intervention. These questions attempted to differentiate between thinking skills lessons and other lessons. It was only by implementing the approach it was realised that this contradicted the principle of infusion and transfer as espoused in the ACTS intervention. The other concern with using a self-

report instrument is that as children grow and develop, their ability to express views and opinions improves. Such a factor may mean differences exist between the *experimental* and *waiting list control* group responses on the basis of age and developmental progression alone.

The procedure for administration of the student measure was conducted by individual class teachers at a time of their choosing at the end of the summer term. This may mean the teacher primed pupils into a certain set of responses linked to their own perspective. This may have had a positive or negative influence according to the class teacher concerned. This procedure also meant the actual time of administration was not consistent across classes. Such a factor may have influenced results according to the time of day, lesson preceding, proximity of playtime and day of the week. These confounding variables were not accounted for, or controlled in a scientific manner. Future studies would need to standardise the procedure of administration for this measure.

The analysis of pupil results was conducted by the experimenters involved, which may have introduced bias into proceedings. It could be that researchers had predetermined ideas of what they wanted to find and located only data that was consistent with their original hypotheses. Furthermore pupil data was analysed in close proximity to staff data, which may have influenced themes constructed. Furthermore, no reflexivity checks with the pupils concerned occurred which limits the accuracy of interpretation. Instead

teacher data and colleague consultation provide a validity check on data analysis.

The qualitative phase of the project could be improved upon in several ways to extend both its reliability and validity. Firstly questionnaire data could be triangulated through the use of focus groups or semi-structured interviews alongside using SWOT (Strengths, Weaknesses, Opportunities and Threats) analyses (McKinstery & Topping, 2003). Data analysis could utilise more rigorous methodology such as discourse analysis (Potter & Wetherell, 1987). Finally reflexivity checks could have occurred more regularly with all participants (Stevenson & Cooper, 1997).

Analysis and Interpretation

The study offered the first evaluation of an infusion intervention combining both quantitative and qualitative methodologies. Several limitations in both the analysis and interpretation of such data, however, restrict conclusions drawn. Firstly with regards to the two groups used, no detail is provided as to statistical adjustments made to account for baseline differences. Similarly insufficient information is provided as to sample size. In future the power of the study to detect a statistically significant effect would be needed (Greenhalgh, 2001).

In relation to quantitative data, variables were not clearly operationalised nor adequately managed in analysis. The statistical techniques were justified, but analysis of the data was confined to certain techniques given the complexity

of the research design and the number of variables being analysed. As a result MANOVA's remained the principal method of analysis and simpler procedures were employed on the basis of this i.e. one-way ANOVA's, and t-tests. This may mean statistical interpretation was skewed by the original technique. The quantitative write up also failed to provide detail pertaining to outliers in relation to theory or statistical meaning. Moreover the distinctions between correlation and regression were not distinguished and assumptions made about the nature and direction of causality. In future more detailed information and rationalised assumptions would be required.

The qualitative analysis and interpretation lacked scientific rigour. Firstly the frequencies of responses were not reported on in the thematic analysis. This meant that, although, the quality of the evidence was illustrated the quantity was not. This meant it was difficult to ascertain whether the themes were expressed by all participants or a subgroup alone. Another issue in the analysis was the role of researchers in acting as the initial point of both collection and interpretation of the data. This may have meant responses were skewed by what researchers hoped to find. Similarly reflexivity checks were not sufficiently detailed so as to allow for replication. In addition they occurred with EP colleagues first followed by participants both of whom could have been influenced by the involvement of the researcher. Improvements would be needed in this area to ensure reflexivity with all participants occurred without the impact of the researcher. Likewise researchers would need to be blind as to participant responses from the group allocated in the coding and interpretation procedure (Greenhalgh, 2001).

Conclusion

The present study develops the empirical evidence base in relation to the infusion methodology for teaching thinking skills. Such cognitive instruction heralds considerable developments for psychology and education alike (DfEE, 1999; Higgins et al., 2004; TES, 2002; Watkins et al., 2001). In particular the potential to qualitatively enhance children's thinking to a higher level has huge advantages to every sphere of society - from students and schools to industry and economies. This research takes one of the first steps in the realms of Educational psychology to ascertain if a programme can achieve such positive change. It is now up to us as a profession to decide if the thinking skills revolution is a worthy cause and should be accepted as a fundamental fact for education or dismissed as futile faddism. We have the potential to determine what direction the future may take and should welcome it with open minds.

References

Adey, P. (2002). *Teacher Empowerment.* Paper presented at the 10th International Conference for Thinking. Harrogate, June 14th –16th, 2002.

Adey, P. (1992). The CASE results: implications for science teaching. *International Journal for Science Education,* 14, 137-146.

Adey, P., Robertson, A. and Venville, G. (2002). Effects of a cognitive acceleration programme on Year 1 pupils. *British Journal of Educational Psychology*, 72, 1-25.

Adey, P. and Shayer, M. (1993). An exploration of long term far-transfer effects following an extended intervention program in the high school curriculum. *Cognition and Instruction,* 11 (1), 1-29.

Adey, P. and Shayer, M. (1994). *Really Raising Standards: Cognitive Intervention and academic achievement.* London: Routledge.

Allinson, J. and Hayes, C. (1988). The Learning Style Questionnaire: an alternative to Kolb's Inventory? *British Journal of Management*, 25, 269-81.

Allinson, J. and Hayes, C. (1990). Validity of the Learning Styles Questionnaire. *Psychological Reports*, 67, 859-866.

Alvino, J. (1990). A Glossary of Thinking-Skills Terms. *Learning,* 18 (6), 50.

Anderson, M. (1992). *Intelligence and development; A cognitive theory.* London; Blackwell.

Ames, C. (1986). Effective Motivation; The contribution of the learning environment. In R.S. Feldman (Ed), *The Social Psychology of Education*, (p.235-256). Cambridge; Cambridge University Press.

Aronson, J. (1994). A pragmatic view of Thematic Analysis. *The Qualitative Report*, 2 (1), 1994 at www.nova.edu/sss/QR/BackIssues/QR2-1/aronson.html.

Atkinson, S. (1998). Cognitive style in the context of design and technology project work. *Educational Psychology,* 18 (2), 183-194.

Baird, J. R. (1986). Improving learning through enhanced metacognition; a classroom study. *European Journal of Science Education*, 8, 263-282.

Banister, P. (1994). Report Writing. In P. Banister, E. Burman, I. Parker, M. Taylor and C. Tindall. *Qualitative Methods in Psychology.* Buckingham: Open University Press.

Barker,C. and Pistrang, N. (2005). Quality Criteria Under Methodological Pluralism: Implications for Conducting and Evaluating Research. *American Journal of Community Psychology*, 35, (3/4), 201- 212.

Baxter, J. (2002). Discussion paper – Creative Partnerships; public and voluntary services working together. *Educational Psychology in Practice*, 18 (1), 63-72.

Baxter ,J. and Frederickson, N. (2005). Every Child Matters: Can educational psychology contribute to radical reform? *Educational Psychology in Practice*, 21(2), 87-102.

Belsky, J. and Fearon, R. M. P. (2002). Early attachment security, subsequent maternal sensitivity, and later child development: Does continuity in development depend upon continuity of care giving? *Attachment and Human Development*, 4 (3), 261-387.

Biggs, J. (1987). *The Learning Process Questionnaire (LPQ): Manual.* Hawthorn, Vic: Australian Council for Educational Research.

Biggs, J. and Moore, P. (1993). *The Process of Learning.* Third Edition. New York: Prentice Hall.

Binsted, D. (1982). Design for learning in management training and development: A View. *Learning and Management Training*, 4 (8), 1-31.

Birmingham City Council. (1997). *Signposts.* Windsor. NFER- Nelson.

Blagg, N. (1991). *Can we teach intelligence?* Hillsdale, NJ: Lawrence Erbaum Associates.

Blank, L.M. (2000). A metacognitive learning cycle: a better warranty for student understanding. *Science Education,* 84, 486-506.

Bond, T. (2002). *Changing Minds and Empirical Evidence: Can we measure change?* Paper presented at the 10th International Conference of Thinking, Harrogate, June 14-16, 2002.

Brace, N., Kemp, R. and Snelgar, R. (2003). *SPSS for Psychologists. A Guide to Data Analysis using SPSS for Windows.* Second Edition. Palgrave Macmillan.

British Psychological Society. (2000). *Code of Conduct, Ethical Principles and Guidelines.* Leicester: BPS.

Brown, A.L. (1987). Metacognition, executive control, self-regulation and other more mysterious mechanisms. In R. Kluwe and F. Weinert (Eds.), *Metacognition, motivation and understanding* (p.65- 116). London: Lawrence Erlbaum.

Burden, R.L. (1998) Assessing children's perceptions of themselves as learners and problem solvers. *School Psychology International*, 19(4), 291-305.

Burden, R.L. and Fraser, B.J. (1993). Use of classroom environment assessments in school psychology: a British perspective. *Psychology in the Schools,* 30, 232-240.

Burden, R. and Nichols, L. (2000). Evaluating the process of introducing a thinking skills programme into the secondary school curriculum. *Research Papers in Education,* 15, 293-306.

Burnett, P. C. (1994). Self-concept and self-esteem in elementary school children. *Psychology in the Schools*, 11, 164-171.

Burnett, P. C. (1998). Measuring behavioural indicators of self-esteem in the classroom. *Journal of Humanistic Education and Development*, 37,107-116.

Burnett, P. C. (1999). Children's self talk and academic self-concepts: the impact of teachers' statements. *Educational Psychology in Practice*, 15, 195 - 200.

Burnett, P. C. and Proctor, R. M. (2002). Elementary school students' learner self-concept, academic self-concepts and approaches to learning. *Educational Psychology in Practice*, 18 (4), 325- 333.

Busato, V. V., Prins, J., Elshout, J. J. and Hamaker, C. (1998). Learning styles: a cross-sectional and longitudinal study in higher education. *British Journal of Educational Psychology,* 68 (3), 427-441.

Calderhead, J. (1996). Teachers' beliefs and knowledge (p.709-725) in D. C. Berliner and R. C. Calfee (Eds), *Handbook of Educational Psychology*. New York: Macmillan.

Cardelle-Elawar, M. (1992). Effects of teaching metacognitive skills to students with low mathematics ability. *Teaching and Teacher Education*, 8, 109-121.

Cardelle-Elawar, M. (1995). Effects of metacognitive instruction on low achievers in mathematics problems. *Teaching and Teacher Education*, 11, 81-95.

Case. R. (1978). A developmentally based theory and technology of instruction. *Review of Educational Research*, 48 (3), 439 – 463.

Chi, M, T, H. (1996). Constructing self explanations and scaffolded explanations in tutoring. *Applied Cognitive Psychology*, 10, 33-49.

Coffman, W. E. (1980). Those achievement tests – How useful? *Executive Review*, 1 (1), 2 – 5.

Coles, M.J. (1993). Teaching Thinking: principles, problems and programmes. *Educational Psychology*, 13 (3-4), 333 – 344.

Coles, M.J and Robinson, W.D. (Eds) (1991). *Teaching Thinking: a survey of programmes in education*. London: Bristol Classical Press.

Connell, J.P. (1985). A new multidimensional measure of children's perceptions of control. *Child Development*, 56, 1018-1041.

Connor, M J. (2002). *The Development of Thinking Skills.* Surrey Educational Psychology Service, April 2002.

Connor, M.J. (1997). *Critical Appraisal.* Surrey Educational Psychology Service, November 1997.

Costa, A. L. (2001). *Developing Minds: A resource book for teaching thinking.* Third Edition. Alexandria VA: Association for supervision and curriculum development.

Cotton, K. (1991). *Teaching Thinking Skills.* School Improvement Research Series (SIRS), North West Regional Educational Laboratory. Available at hhtp://www.nwrel. org/scpd/sirs/6/cu11.html.

Crooks, C. K. (1994). *Computers and the collaborative experience of learning.* London: Routledge.

Crooks, C. K. (1995).On resourcing a concern for collaboration within peer interactions. *Cognition and Instruction*, 13(4), 541-547.

Crooks, C. K. (1998). Computer networks in education. *The Psychologist*, 11(8), 378-380.

Crossley, M. L. (2000). *Introducing narrative psychology: Self, trauma and the construction of meaning.* Buckingham: Open University Press.

Dart, B., Pillay, H. and Burnett, P. C. (2000a). Australian and Filipino students' approaches to learning, conceptions of learning and learner self-concepts; a cross cultural comparison. *Educational Research Journal*, 15, 143 – 165.

Dart, B., Burnett. P. C., Purdie. , N., Boulton- Lewis, G., Campbell, J. and Smith, D. (2000b). Influences of students' conceptions of learning and the classroom environment on approaches to learning. *Journal of Educational Research*, 93, 262 – 270.

Dasen, P. R. (1972). Cross cultural Piagetian research: A summary. *Journal of Cross Cultural Psychology,* 3 (1), 23-40.

De Bono, E. (1992). *Teach your child to think.* London: Penguin.

De Corte, E. (1990). Towards powerful learning environments for the acquisition of problem solving skills. *European Journal of Psychology of Education*, 5 (1), 5-19.

De Corte, E. (2002). Creating *high-powered learning communities for thinking and problem solving.* Paper presented at the International Conference for Thinking, Harrogate, June 14 -16th, 2002.

De Corte, E., Verschaffel, .L. and De Ven, A.V. (2001). Improving text comprehension startegies in upper primary school children: a design experiment. *British Journal of Educational Psychology*, 71, 531-559.

Department for Education and Employment. (1999). *From Thinking Skills to Thinking Classrooms; A review and evaluation of approaches for developing pupils thinking.* Research Report RR115. London, DfEE.

Department for Education and Employment. (2000). *Educational Psychology Services (England): Current Role, Good Practice and Future Directions.* The Research Report. London, DfEE.

Department for Education and Science (DES). (1989). *Discipline in Schools.* Report of the Committee of Enquiry Chaired by Lord Elton. London: HMSO.

Department for Education and Skills. (2002). *ImpaCT2: the impact of information and communication technologies on pupil learning and attainment.* London: Department for Education and Skills.

Department for Education and Skills. (2004). *Five Year Strategy for Children and Learners. Putting people at the heart of public services.* London: The Stationery Office (TSO).

Dillenbourg, P. (ed) (1999). *Collaborative Learning: cognitive and computational approaches.* Amsterdam; Pergamon.

Dodge, K. A., McClaskey, C. L. and Feldman, E. (1985). Situational approach to the assessment of social competence in children. *Journal of Consulting and Clinical Psychology,* 53(3), 344-353.

Dodge, K.A., Coie, J.D. and Brakke, N.P. (1982). Behaviour patterns of socially rejected and neglected preadolescents; the roles of social approach and aggression. *Journal of Abnormal Child Psychology*, 10 (3), 389-410.

Doheer, L., Reynolds., S., Wetherly,. J., and Evans, E.H. (2005). Young children's ability to engage in cognitive therapy tasks: Associations with age and educational experience. *Behavioural and Cognitive Psychotherapy*, 33(2), 201-215.

Donaldson, M. (1978). *Children's Minds.* Glasgow: Fontana.

Drew, P. (1995). Conversation analysis. In Smith J.A., Harre, R. and Van Langenhove, I. (1995) *Rethinking methods in psychology.* London: Sage.

Dweck, C.S., and Bempechat, J. (1993). Children's theories of intelligence: Consequences for learning. In Paris, S.G., Olson, G.M. and Stevenson, H.W. (Eds). *Learning and motivation in the classroom.* (p.239-256).Hillsdale, NJ: Lawrence Erlbaum.

Educational Testing Service.(1993).*Tasks in Critical Thinking.* PO Box 6000, Princeton, NJ 08541.

Elliott, R., Fischer, C.T. and Rennie, D.L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38, 215-229.

Ennis, R.H., Gardiner, W., Gazzetta, J., Morrow, R. Paulus, D. and Ringel, L. (1964). *Cornell Class Reasoning Test.* Illinois Critical Thinking Project, University of Illinois, 1310 S, 6th Street, Champaign, IL 61820.

Ennis, R.H., Gardiner, W., Gazzetta, J., Morrow, R. Paulus, D. and Ringel, L. (1964). *Cornell Conditional Reasoning Test.* Illinois Critical Thinking Project, University of Illinois, 1310 S, 6th Street, Champaign, IL 61820.

Erlandson, D. A., Harris, E. L., Skipper, B. L. and Allen, S.D. (1993). *Doing Naturalistic Inquiry.* London: Sage.

Facione, P. (1990). *The Californian Critical Thinking Skills Test.* The California Academic Press, 217 La Cruz Avenue, Millbrae, CA 94030.

Facione, P. (1998). *The Test of Everyday Reasoning.* Derived from the Californian Critical Thinking Skills Test. The California Academic Press, 217 La Cruz Avenue, Millbrae, CA 94030.

Feuerstein, R., Y Hoffman, M.B. and Miller, R. (1980). *Instrumental Enrichment: An intervention for cognitive modifiability.* Baltimore, MD: University Park Press.

Field, A. (2003). *Discovering statistics using SPSS for Windows: Advanced Techniques for Beginners. (Introducing Statistical Methods).* Sage Publications Ltd.

Fisher, A. (2002). *Assessing Thinking.* Paper presented at the International Conference of Thinking, Harrogate, June 14-16th, 2002.

Fitz-Gibbon, C.T. (2002). *What do schools need? How can Educational Psychologists contribute?* Keynote address at the Association of Educational Psychologists Annual Conference, Hinckley, October 24-26th, 2002.

Fox, M. and Sigston, A. (1992). Connecting Organisational Psychology, Schools and Educational Psychologists in Wolfendale, S.; Bryans, T.; Fox, M., Labram, A., and Sigston, A. (Eds).*The Profession and Practice of Educational Psychology*. London: Cassell.

Fraser, B.J. (1979). *Test of Enquiry Skills*. Australian Council for educational Research Limited, Frederick Street, Hawthorn, Victoria 31222, Australia.

Fraser, B.J., Malone, J. A. and Neale, J. M. (1989). Assessing and improving the psychosocial environment of mathematics classrooms. *Journal of Research in Mathematics Education,* 20, 191-201.

Fuchs, L .S. Fuchs, D., Kazdan, S. and Allen, S. (1999). Effects of peer-assisted learning strategies in reading with and without training in elaborated help giving. *Elementary School Journal*, 99, 201-220.

Galloway, D. and Rogers, C. (1994) Motivational style: a link in the relationship between school effectiveness and child behaviour. *Educational and Child Psychology,* 11 (1), 16-25.

Gardner, H. (1993). *Multiple Intelligences: The theory in practice*. New York: Basic Books.

Georghiades, P. (2000). Beyond conceptual change learning in science education: focusing on transfer. *Educational Research*, 42, 119-139.

Goos, M. (1999). *Development and use of the Metacognitive Self-Knowledge Questionnaire.* Unpublished manuscript, Graduate School of Education, The University of Queensland, Australia.

Gorodetsky, M., Barak J., Kaur, R, Xamir J and Tirosh, R. (2002). *A Knowledge Constructing School.* Paper presented at the International Conference of Thinking, Harrogate, June 14-16th, 2002.

Greenhalgh, T. (2001). *How to read a paper. The basics of evidence based medicine.* BMJ Books: London.

Hagley, F. (1987). *Suffolk Reading Scale*. Windsor: NFER – Nelson.

Hatcher, R. (2002). *The Business of Education*. Keynote address at the Association of Educational Psychologists Annual Conference, Hinckley, October 24-26th, 2002.

Higgins, S. (2002). *Teaching Thinking through pedagogical strategies in primary schools.* Paper presented at the International Conference of Thinking, Harrogate, June 14-16th, 2002.

Higgins, S., Baumfield, V., Lin., M., Moseley.,D., Butterworth., M., Downey., G., Gregson., M., Oberski.,I., Rockett.,M. and Thacker.D. (2004). *Thinking skills approaches to effective teaching and learning: what is the evidence for impact on learners.* In: Research Evidence in Education Library. London: EPPI-Centre, Social Science Research Unit, Institute of Education.

Hill, C .E., Thompson, B. J., and Williams, E. N. (1997). A guide to conducting consensual qualitative research. *The Counselling Psychologist*, 25, 517- 572.

Honey, P. and Mumford, A. (1982). *The Manual of Learning Styles*. Revised version. Maidenhead: Peter Honey.

Honey, P. and Mumford, A. (1986). *Using your learning styles*. Maidenhead: Peter Honey.

Hughes, G. (1998). *Critical Thinking Interview*. Minnesota State Colleges and Universities. 141 Warwick Street, S.E., Mpls., MN 55414.

Inhelder, B. and Piaget, J. (1958). *The growth of logical thinking*. London: Routledge and Kegan Paul.

Joiner, R., Littleton, K., Faulkner, D. and Miell, D. (2000). *Rethinking collaborative learning*. London: Free Association Press.

Jenner, S. and Gravenstede, A. (1998). Effective Discipline, Effective Schools. *Educational Psychology in Practice*, 14 (1), 60-65.

Jensen, A., Malcolm, L., Phelps. , F. and Stoker, R. (2002). Changing patterns of thinking; individuals and organisations. *Educational Psychology in Practice*, 18 (1), 35 – 45.

Kail. , R. and Bisanz., J. (1992). The information processing perspective on cognitive development in childhood and adolescence. In R.J. Sternberg and C.A. Berg (Eds*) Intellectual Development*, p 229-260. Cambridge University Press.

Kearney, D. (1994). *Coaching*. Yeovil, Somerset: Organisational and Personal Development Consultants.

Kember, D., Biggs, J. and Leung. D. Y. P. (2004). Examining the multidimensionality of approaches to learning through the development of a revised version of the Learning Process Questionnaire. *British Journal of Educational Psychology*, 74 (2), 261-279.

Knowles, M. (1984). *The Adult Learner: A Neglected Species*. Third Edition. Houston: Gulf Publishing Co.

Kolb, D.A. (1976 and 1985). *Learning Style Inventory and Technical Manual.* Englewood Cliffs, NJ: Prentice Hall.

Kolb, D.A., Rubin, R. M. and McIntyre, J. M. (Eds). (1984).*Organizational Psychology- Readings on Human Behaviour in Organisations.* New Jersey: Prentice Hall.

Lazar-Morrison, C. (1980). *A review of the literature on test use.* Los Angeles: Centre for the study of Education. ERIC Document Reproduction. No. ED. 204 411.

Leininger, M. M. (1985). *Qualitative Research Methods in Nursing.* Orlando, FL: Grune and Stratton.

Light, P. and Littleton, K. (eds). (1999). *Social processes in children's learning.* Cambridge: Cambridge University Press.

Lipman, M. (1991). *Thinking In Education.* Cambridge; Cambridge University Press.

Lipman, M., Sharp. A. M. and Oscanyan, F.S. (1980). *Philosophy in the classroom.* Philadelphia: Temple University Press.

Littleton, K. and Light, P.(eds). (1999). *Learning with computers: analysing productive interaction.* London: Routledge.

Lohman, D.F., Thorndike, R. L. and Hagen, E. P. (2001). *Cognitive Abilities Test.* Third edition. Windsor. NFER- Nelson.

Mackay, T. (1999). *Quality Assurance in Education Authority Psychological Services: a self-evaluation using performance indicators.* Edinburgh: Scottish Executive Education Department.

MacKay, T. (2002). Discussion Paper- The future of Educational Psychology. *Educational Psychology in Practice*, 18 (3), 245-253.

Meadows, S. (1998). Children learning to think: learning from others? Vygotskian theory and educational psychology. *Educational and Child Psychology,* 15(2), 6-13.

Mercer, N. (1995). *The guided construction of knowledge.* Clevedon, Avon: Multilinigual Matters.

Mercer, N., Wegerif, R., and Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal,* 25(1), 95-111.

McGuinness, C. (1993) Teaching Thinking: New Signs for Theories of cognition. *Educational Psychology*, 13 (3-4), 305-316.

McGuinness, C. (1999). *From Thinking Skills to Thinking Classroom: a review and evaluation of approaches for developing pupil's thinking.* DfEE Research Report, RR115.

McGuinness, C. (2000). *ACTS; A methodology for enhancing thinking skills across the curriculum.* C.mcguinness@qub.ac.uk.

McGuinness, C. (2002). *Metacognition in Classrooms.* Paper presented at the 10th International Conference of Thinking, Harrogate, June 14 -16th, 2002.

McGuinness, C. (2003). *Developing Thinking Skills across the Curriculum; ACTS in Northern Ireland.* Keynote Address presented at the Association of Educational Psychologists Annual Conference, Belfast, 31st October, 2003.

McGuinness, C., Curry, C., Greer, B., Daly, P., and Salters, M. (1997). *Final Report on the ACTS project. Phase 2.* Belfast: Northern Ireland Council for Curriculum Examinations and Assessments.

McGuinness, C. and Nisbet, J. (1991). Teaching Thinking in Europe. *British Journal of Educational Psychology*, 61, 174-186.

McKee, W.T. and Witt, J. C. (1990). Effective teaching: a review of instructional and environmental variables. In Gutkin, T.B. and Reynolds, C. R. (eds). *The Handbook of School Psychology.* New York: Wiley.

McKinstery, J., and Topping, K. J. (2003). Cross- age Peer Tutoring of Thinking Skills in the High School. *Educational Psychology in Practice*, 19(3), 199-217.

Morton, J. and Frith, U. (1995). Causal modelling; a structural approach to developmental psychopathology. In Cicchetti, D. and Cohen, D.J. (Eds). *Manual of Developmental psychopathology.* New York; Wiley.

Munro, J. (1999). Learning more about Learning Improves Teacher Effectiveness. *School Effectiveness and School Improvement*, 10 (2), 151-171.

Nangle, D. W., Ellis, J.T. and Hansen, D .J. (1994). Clinical utility of the Taxonomy of Problematic Social situations for Children (TOPS): temporal stability and convergence with peer measures. *Education and Treatment of Children*, 17(1), 129-138.

Neale, M. D. (1989). *The Neale Analysis of Reading Ability – Revised.* Windsor: NFER- Nelson.

Nisbet, J. (1993). The Thinking Curriculum. *Educational Psychology*, 13 (3-4), 281-290.

Norris, S. P. and King, R. (1983). *Test on Appraising Observations.* Institute for Educational Research and Development, Memorial University of Newfoundland, St John's, Newfoundland, Canada, A1B 3X8.

Norwich, B. (1998). Developing a Self-assessment of Learning Inventory for use in action planning. *Educational Psychology in Practice*, 14 (3), 147- 155.

Parker, I. (1992). *Discourse Dynamics; Critical analysis for social and individual psychology.* London: Routledge.

Parker, I. (1994). Discourse Analysis. In P. Banister, E. Burman, I. Parker, M. Taylor and C. Tindall. *Qualitative Methods in Psychology.* Buckingham: Open University Press.

Perkins, D. N. and Grotzer, T. A. (1997). Teaching intelligence. *American Psychologist,* 52 (10), 1125-1153.

Perkins, D. N., Jay, E. and Tishman, S. (1993). Beyond abilities: A dispositional theory of thinking. *The Merill-Palmer Quarterly*, 39, 1-21.

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development,* 5, 1-12.

Pithers, R. T. and Soden, R. (2000). Critical Thinking in Education: A Review. *Educational Research*, 42 (3), 237-249

Portwood, M. (2002). *A metabolic basis for dyslexia, dyspraxia and attention deficit hyperactivity disorder.* Keynote Address at the Association of Educational Psychologists Annual Conference, Hinckley, October 24-26th, 2002.

Potter, J. and Wetherell, M. (1987). *Discourse and social psychology; Beyond attitudes and behaviour.* London: Sage.

Pramling, I. (1988) Developing children's thinking about their own learning. *British Journal of Educational Psychology*, 58, 266-278.

Proctor, B. and Inskipp, F. (1991). Group Supervision. Chapter 6 in Hawkins, P. and Shohet, R. (2000). *Supervision in the helping professions.* Buckingham: Open University Press.

Riding, R.J. (1991). *Cognitive Styles Analysis.* Birmingham: Learning and Training Technology.

Riding, R.J., Burton, D. Rees, G. and Sharrat, M. (1995). Cognitive Style and Personality in 12-year-old children. *British Journal of Educational Psychology*, 65, 113-124.

Ritchie, S. M. and Edwards, J. (1996). Creative thinking instruction for aboriginal children. *Learning and Instruction,* 6, 59-75.

Ross, J.D. and Ross, C.M. (1976). *Ross Test of Higher Cognitive Processes (RTHCP).* Academic Therapy Publications, 20 Commercial Blvd, Novato, CA 94947.

Rudestam, K.E. and Newton, R.R. (1992). *Surviving your Dissertation.* Newbury Park, California: Sage.

Sackett, D. L., Haynes, R. B., Guyatt, G. H., and Tugwell, P. (1991). *Clinical epidemiology- a basic science for clinical medicine*. London: Little, Brown.

Salomon, G. and Perkins, D.N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Journal of Educational Psychology*, 24, 113-142.

Savell, J.M., Twohig, P.T. and Rachford, D.L. (1986). Empirical Status of Feuerstein's 'Instrumental Enrichment' (FIE) technique as a method of teaching thinking skills. *Review of Educational Research*, 56, 381-409.

Senge, P. (1993). *The Fifth Discipline*. London: Century House.

Shayer, M. and Adey, P. (1993). Accelerating the development of Formal Thinking in Middle and High School Students IV: Three years after a two-year intervention. *Journal of Research in Science Teaching*, 30(4), 351-366.

Shayer, M. and Beasley, F. (1987). Does Instrumental Enrichment work? *British Educational Research Journal*, 13, 101-119.

Sherrard, C. (1997). Qualitative research. *The Psychologist*, 10, 161-162.

Simmons, D., Fuchs, D., Fuchs, L. S., Pate, J. and Mathes, P. (1994). Importance of instructional complexity and role reciprocity to class wide peer tutoring. *Learning Disabilities Research and Practice*, 9, 203-212.

Simon., H., A. (1979). Information processing models of cognition. *Annual Review of Psychology*, 30, 363-396.

Smith, A. (1996). *Accelerated learning in the classroom.* Staffordshire; Network Educational Press.

Smith, J.A. (1996). Beyond the divide between cognition and discourse: using interpretative phenomenological analysis in health psychology. *Psychology and Health,* 11, 261-271.

Splitter, L. J. and Sharp, A. (1995). *Teaching for better thinking: The classroom community of inquiry.* Melbourne: Australian Council for Educational Research.

Sternberg, R. J., and Bhana, K. (1986). Synthesis of research on the effectiveness of intellectual skills programs: Snake oil remedies or miracle cures? *Educational Leadership*, 44(2), 60 -67.

Stevenson, C. and Cooper, N. (1997). Qualitative and quantitative research. *The Psychologist,* 10, 159 – 160.

Stewart, L. and Smardon, D. (2002). *Teachers and Learners participating in Critical Thinking for effective change.* Paper presented at the 10th International Conference of Thinking, Harrogate, June 14th –16th, 2002.

Stoiber, K.C. and Kratochwill, T.R. (2000). Empirically supported interventions and School Psychology: conceptual and practical issues. *School Psychology Quarterly*, 15, 75-105.

Swartz, R. and Parks, S. (1994). *Infusing the teaching of critical and creative thinking into content instruction: A lesson design handbook for the elementary grades.* California: Critical Thinking Press and Software.

Taylor, S.J., and Bogdon, R. (1984). *Introduction to qualitative research methods: The search for meanings.* New York: John Wiley and Sons.

TES (2002). Thinking: the next big idea....*Times Educational Supplement.* June 14th 2002.

Thorndike, R.L. and Hagen, E. (1986). *Cognitive Abilities Test.* Second Edition. Windsor. NFER- NELSON.

Topping, K. (2002). Paired Thinking. Developing Thinking Skills through Structured Interaction with Peers, parents and volunteers. In Reid, G. and Wearmouth, J. (ed) *Dyslexia and Literacy: Theory and Practice.* Chapter 18. John Wiley and Sons Ltd. 2002.

Topping, K. (2003). BP Read On Project. Paired Reading and Thinking. Extending into Paired thinking. At www.dundee.ac.uk/psychology/readon/welcome.htm

Topping, K. J. and Bryce, A. (2004). Cross-age peer tutoring of reading and thinking in the primary school; a controlled study of influence on thinking skills. At www.dundee.ac.uk/psychology/readon/welcome.htm

Trickey, S. (2004). Psychologists making a difference to all children. Promoting Learning and Thinking through Philosophical Enquiry. *Educational Psychology in Scotland*, 6 (3), 24- 28.

Vye, N.J., Goldman, S. R., Voss., J. F., Hmelo, C. , and Williams, S. (1997). Complex mathematical problem solving by individuals and dyads. *Cognition and Instruction,* 15(4), 435-484.

Vygotsky, L.S. (1978). *Mind in Society; the development of higher order psychological processes.* Edited and translated by Cole, M., John-Steiner, V., Scribner, S., Souberman, E. Cambridge, MA, Harvard University Press.

Watkins, C. (2003). *Thinking and Learning skills- what next?* Presentation at The Institute of Education. February 7th, 2003.

Watkins, C., Carnell, E., Lodge, C., Wagner, P., and Whalley C. (2001). Learning about Learning enhances performance. *The National School Improvement Networks Bulletin,* Spring 2001.

Watson, G. and Glaser, E.M. (1980). *Watson-Glaser Critical Thinking Appraisal.* The Psychological Corporation, 555 Academic Court, San Antonio TX 78204.

Watt, S. (2003). *Qualitative Research Methods: A Brief Introduction.* Lecture and workshop at UCL on D.Ed. Psy, October 2003.

Wegerif, R., Higgins, S., Mercer, N., Dawes, L. and Sams, C. (2002).*Thinking Together: Research into Practice.* Paper presented at the 10th International Conference of Thinking. June 14th-16th, 2002.

Wegerif, R., Mercer, N., and Dawes, L. (1999). From social interaction to individual reasoning: An empirical investigation of a possible socio-cultural model of cognitive development. *Learning and Instruction*, 9 (6), 493-516.

Wegerif, R. and Scrimshaw, P. (eds). (1997). *Computers and talk in the primary classroom.* Clevedon: Multilingual Matters.

White, B.Y. and Frederiksen, J.R. (1998). Inquiry, modelling and metacognition: Making science accessible to all students. *Cognition and Instruction*, 16(1), 3-118.

Williams, G., Whittome, R. and Watts, P. (2003). PASS: Measuring attitudes to learning. *Learning and Skills Research*, 6 (2), 36.

Willig, C. (2001). *Qualitative research in psychology; A practical guide to theory and method.* Buckingham: Open University Press.

Wilson, D. K. (1986). An investigation of the properties of Kolb's Learning Style Inventory. *Lodi*, 7 (3), 1-15.

Wilson, V. (2000). *Can Thinking Skills be taught? A paper for discussion.* Scottish Council for Research in Education.

Yin, R.K. (1994). *Case Study Research.* London; Sage.

Ysseldyke. J. E. and Christenson, S. L. (1987). *TIES: The Instructional Environment Scale.* Austin, Texas: Pro-Ed.

Ysseldyke. J. E. and Christenson, S. L. (1993). *TIES-II: The Instructional Environment System II. A System to Identify a Student's Instructional Needs.* Longmont, CO: Sopris West.

Zimmerman, B.J. (1998). Academic studying and the development of a personal skill: self regulatory perspective. *Educational Psychologist*, 33 (2/3), 73-76.

Appendices

Appendix 1: Preparatory Phase

The preparatory phase of the research methodology was conducted during April – September 2002. The purpose and aim for this phase will be outlined alongside the rationale. Methodological considerations will be illustrated within the process of consultation. Finally a conclusion as to the main implications for the study will be highlighted and discussion of the preparatory phase made.

Purpose of Preparatory Phase

The purpose of the preparatory phase is two-fold in informing both the theoretical and practical underpinnings of the research methodology. The principal objective is to ascertain *how* to quantify differences in students over time and over *what* dimensions as a result of a thinking skills intervention. In particular the following aims are to be addressed:-

Theoretical Aims

1. To sharpen the focus of the study in relation to manageable and meaningful data collection.
2. To inform the selection of measurement devices to investigate and quantify variables identified in the five hypotheses from a qualitative and quantitative perspective.

Practical Aims

3. To clarify the time required to conduct assessments and the procedures of administration.
4. To identify any issues hereto undetected.
5. To develop prompt materials and a rationale for participants.

The rationale behind using a preparatory phase instead of a pilot study was manifold. Firstly it was not considered ethically sound to put participants through assessment procedures when there would be no value added impact or relevance for them, as in a pilot study. This was decided in consultation with schools and staff. Additionally it was apparent that information and decisions could be made through consultation with leading professionals in the area and evaluation of previous studies. Furthermore, measurement devices could be obtained to examine potential, alongside liaison with experts who had experience of usage. Finally in terms of time, restrictions from standard fund finance meant a project had to be implemented by September 2002. This meant a preparatory phase occurred as opposed to a potentially more lengthy pilot study.

The overriding aim of the preparatory phase is to overcome the limitations exposed in previous research. In particular the application of psychology is to be used to address previous research flaws from the lack of control groups to the use of inappropriate measures. Similarly unresolved issues will be targeted in terms of both time and optimum age (Blagg, 1991; Coles, 1999; Nisbet, 1993). The restricted size and scope of earlier work will likewise be

improved (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000). In this manner, research in the area of thinking skills will be improved by using a more scientifically rigorous methodology. Finally the preparatory phase will evidence the originality of the research in designing the first long term evaluation of an infusion thinking skills intervention with an appropriate control. The infusion approach, namely, ACTS (Activating Children's Thinking Skills) was selected as there was a need for long-term quantitative evaluation in this area (McGuinness, 1999). Moreover, the intervention accommodates for the issues of transfer and maintenance which were areas omitted in earlier research (DfEE, 1999; McGuinness, 2003). In this fashion the preparatory phase intends to extend research in the domain of thinking skills and develop the psychological theory and practice base.

Methodological Considerations

Decisions regarding the research methodology occurred through a process of consultation on a series of different levels. Initially negotiation with the county LEA took place. This involved securing a standard fund bid linking into the Education Development Plan for Learning and Teaching Strategies. The authority agreed to allow for an intervention evaluation study to occur to inform strategy and future policy. An extensive period of consultation followed with schools to establish current approaches being implemented, optimum year groups and the possibility of real life research in working schools. Links and discussions with external specialists were next made with Professor Carol McGuinness (Queens University Belfast), Dr. Steve Trickey

(Clackmannashire and Dundee University), Dr. Chris Watkins (Institute of Education), Anne Robertson (Hammersmith and Fulham), Dr. Steve Higgins (Newcastle University), Dr. Rupert Wegerif (Open University) and Dr. Keith Topping (Dundee University). These discussions and correspondences permitted an evaluation of the current status of thinking skills research so as to inform the methodology for this study. Liaison, likewise, took place with Dr. Phil Reed and Dr. Simon Watts at University College London regarding appropriate methods of quantitative and qualitative data collection. On this basis a consideration of research at a national level and local level in terms of standard fund requirements were carried out. Such consultation ensured that a thorough and extensive knowledge of the research methodology was achieved, alongside an understanding of its potential application. Key areas are now discussed.

Research Design

Previous research in the area of thinking skills interventions have not utilised control groups to account for the range of confounding variables. Indeed as Sternberg and Bhana (1986) indicate this is a longstanding issue with *"most studies involved inadequate control groups, and some entailed none at all"* (p. 60). In response to this it was decided the study would use a design in which systematic bias could be minimised through the use of an effective control group. An intervention evaluation design was adopted as it would ascertain whether the particular thinking skills intervention of ACTS (Activating Children's Thinking Skills) made a difference between participants from randomly selected groups. Such a design would mean that changes in

participants could not be attributed to the passage of time, individual histories or natural developmental progression.

In particular, four control group designs were evaluated to gauge effectiveness. These were reviewed in terms of scientific rigour, ethical soundness and applicability to the school setting. Discussions occurred at an LEA, UCL and school level to consider these options. A *waiting list control* group was considered most appropriate as it compares an intervention against participants who will subsequently receive the treatment. This controls for the role of expectancy of change and is ethically sound, as participants are not denied a potentially valuable intervention, but rather have it delayed. A *no treatment control*, although, evident throughout thinking skills research does not control for expectancy in this manner. Such a design means a range of extraneous variables would be unaccounted and not controlled for. Similarly it is ethically and educationally a non-viable option as children would receive no input despite considerable assessment. Consultation with both the LEA and schools identified the *waiting list control* as the most manageable and justifiable in the educational environment.

Attention placebo and *alterative intervention controls* were likewise examined, however, at this point the critical aspects of the intervention are unknown with time of intervention unspecified, so neither *attention* nor *alternative intervention controls* could be utilised effectively. The other options to use as an equivalent would also be difficult to identify given the variation in approaches, be it a generic, subject specific or infusion

methodologies. The only possible comparison could have been another infusion approach such as De Bono's (1992) thinking hats. This approach, however, specifies no actual timings or definite allocation of the intervention in the curriculum and so at present would not make a favourable *attention placebo* or *alternative intervention control* group. Similarly the *no treatment, attention placebo or alternative intervention control* options were not considered ethically sound in the Children's Education Service from the school and parental perspectives. Liaison with UCL reiterated these decisions, with a *waiting list control* arguably the most effective and ethical technique to evaluate change. This option also meant resources could be used most effectively in terms of administration and scoring time. Ethical considerations would be met as no student would go through unnecessary testing or be denied an intervention but merely have it delayed. A *waiting list control* would also allow potentially confounding variables to be discredited scientifically.

In terms of the duration of the project, previous research demonstrates that a two-year period of intervention is necessary to have an effect on cognitive change in children (Blagg, 1991; Feuerstein et al., 1980; Shayer & Adey, 1993). The research, therefore, aimed to track participants for at least three years to monitor effects over time, an area essentially ignored in the research to date (DfEE, 1999). This sharpened the focus of the existing studies and would address unresolved issues of transfer as highlighted by numerous professionals (Blagg, 1991; DfEE, 1999; Nisbet, 1993; Watkins et al., 2001; Wilson, 2000). The three-year design meets requirements from others in the

field for if *"it takes a minimum of 2 years to bring about significant improvements in the ability to learn it will only be after that period that improvements in the acquisition of knowledge in basic attainment areas would begin to show"* (Blagg, 1991)(p.38).

Given the omissions in earlier research in evaluating cognitive interventions using traditional pre-post designs (Burden & Nichols, 2000) and quantitative data only, it was decided qualitative evaluations would occur. Such information would add to the understanding of the mechanisms at work in the methodology. This was considered appropriate following discussions with Dr. Steve Trickey at Clackmananshire and Dr. Simon Watts at UCL. It was hoped the use of qualitative data collection would complement and lead to a deeper understanding of the quantitative data.

Measures

Earlier studies lacked reliable and valid measures to tap into children's thinking skills as Blagg (1991) indicated there was *"a limited range of standardised assessment tools suitable for the job"* (p. 39). Furthermore research tended to focus on student academic performance alone with other learner outcomes ignored (Wilson, 2000). This study, therefore, aims to concentrate on the learner holistically; exploring thinking, behaviour, social and affective factors, as areas identified as in need of scientific examination (DfEE, 1999; McGuinness & Nisbet, 1991; Watkins et al., 2001; Wilson, 2000). At the preparatory phase measurement selection was, therefore, informed by addressing previous limitations and extending the scope of

analysis in terms of student factors. In order to address this, the current project aimed to adopt the principle of triangulation to ascertain impact on cognitive, behavioural and emotional factors of students over time. In this manner the piece is original as it explores learner outcomes in a range of domains whereas previous studies failed (DfEE, 1999).

To begin with the assessment of cognitive skills and abilities were reviewed. A range of measurement devices were examined in terms of gauging a child's current thinking and reasoning skills. The possible use of the Ross Test of Higher Cognitive Processes (RTHCP, 1976) was explored. However, this was limited to children aged 10 – 12 years and was demanding for children as it comprised of two hour long sessions. Similarly the lengthy individual administration of either the British Abilities Scales (BAS II) or Wechler Intelligence Scales Competence (WISC III) was considered too time consuming given the study's sample size of approximately 500 children. In addition to this other assessments of thinking were considered. Many of the individual thinking skills assessments focused on one specific area such as The Watson Glaser Critical Thinking Appraisal (Watson & Glaser, 1980) or Californian Critical Thinking Skills Test (CCTST, Facione, 1990). Other devices such as the Tasks in Critical Thinking (Educational Testing Service, 1993), although authentic required trained graders. Numerous measurements existed for the older student at college, as in The Test of Everyday Reasoning (Facione, 1993) and Critical Thinking Interview (Hughes et al., 1998), however, these were not available for primary aged children.

Of the tools available for school aged children, many of the devices were too specific, for example the Cornell Class Reasoning Test (Ennis et al., 1964), which focused on deductive class reasoning alone or the Test on Appraising Observations (Norris & King, 1983), which concentrated on the credibility of statements of observation. Furthermore, assessments were dated with standardisation occurring on a skewed sample population such as the Test of Enquiry Skills (Fraser, 1979) and the Cornell Conditional Reasoning Test (Ennis et al., 1964). Given the limited scope, reliability and standardisation of these measurements, more recently normed and general approaches to assessing thinking and reasoning were explored.

The most time efficient and scientifically valid and reliable instrument for exploring children's thinking was deemed to be the Cognitive Abilities Tests (CATS) (Lohman, Thorndike & Hagen, 1993). This allowed for an exploration of children's thinking and reasoning in verbal, non-verbal and quantitative domains in a scientific and standardised manner. It, therefore, avoided a focus on one area alone and could detect change in a number of different skills. Likewise it was economical in time for administration and scoring as it could be delivered on a whole class basis. Ethical considerations had also been taken into account with the sample on which the measure had been standardised. The more recent norms meant the instrument was more applicable and relevant to the participants under study. Consultation with schools also indicated it would be an appropriate and convenient measure to use. Furthermore, these devices had been successfully used in other recent thinking skills research (Dr. Steve Trickey, Clackmananshire; 2004). The use

of a measure to ascertain changes in thinking and reasoning, targets an area highlighted as in need of examination (Bond, 2003; Watkins et al., 2001). This measure, therefore, offered a view into children's thinking and reasoning skills in a scientific fashion.

The cognitive skills of students were also to be tapped into using attainment measurements in line with other studies (Blagg, 1991; Shayer & Adey, 1993). This was in order to satisfy schools requirements of understanding the impact on academic achievement. Subsequently a range of measures were examined. Individual assessment procedures were explored such as the New Neale Reading Analysis (NARA) (Neale, 1989) and Suffolk Reading Scale (Hagley, 1987). These devices, however, would be too time consuming to individually administer on a sample of approximately 500 students. Furthermore the possible use of group reading or maths assessments (Group Reading Test 5 –14 or Mathematics 5-14) (NfER-Nelson) were rejected on the basis that they offered a restricted perspective of a pupil's progress. Similarly, it was felt ethically unsound to assess children on both the Cognitive Ability Skills Tests and literacy and numeracy given the time and effort required. Schools felt this was inappropriate and as researchers it would to too time consuming to administer and score. As a result it was decided that the study would use measurements of attainment already used by schools. This would ensure consistency across all educational environments. Likewise consultation with schools revealed these measures were of the greatest instructional relevance to themselves as organisations. It was agreed, therefore, that all schools participating in the

study would administer the optional QCA's (Qualification and Curriculum Assessments) at the end of each year in addition to SATS. Both sources of information could then be efficiently used to gauge impact on students' educational attainment. Such a strategy would ascertain transfer of effects into other relevant domains educationally.

These assessments were considered more rigorous than subjective teacher assessment with the possibility of bias, which could skew results. Additionally this procedure would measure attainment over time to establish the potential maintenance of effects. This would address criticisms of earlier thinking skills research in that *"inadequate attention was usually given both to the transfer of training and to durability of training over the long term"* (Blagg, 1991; p. X). In this manner the piece will be original in providing a long-term perspective of an infusion programme's impact accommodating for both transfer and maintenance.

The other area to be explored was the impact on student's behaviour both inside and outside the classroom, a factor hereto ignored in the research (DfEE, 1999). The possible use of observations with students was considered using interval and event sampling techniques. The constraints of time meant this option was not considered viable. Furthermore the snapshot nature of such an observation would be unlikely to reveal in-depth or sufficient data on students. Issues of training observers and inter-rater reliability also compounded difficulties (see Dodge et al., 1982 for details). As a result the standardised measure from Dodge, McClaskey and Feldman

(1985) was examined. This measure, the Taxonomy of Problematic Social Situations (TOPS), explores how students respond to a range of problematic social situations as noted by the teacher. This measure was standardised and allowed for a reliable and valid way of exploring student's behaviours. Equally important, was the fact it was completed by the adult with most contact with the students, who would have a representative view of an individual over contexts and time. Although, this would still be subjective it would be significantly more scientific and replicable compared to an observational approach. Consultation with staff in schools, likewise revealed it was a more relevant tool educationally and more time efficient. The measure is also being used successfully in other ongoing thinking skills research (Trickey, 2004)

Finally the affective components of thinking have been untouched in current research (De Corte, 2002; McGuinness, 1999; Watkins et al., 2001; Wilson, 2000). Consequently a number of measurement devices were reviewed to explore the student's self perceptions and emotions pertaining to themselves as thinkers. Instruments were examined such as the Burnett Self Scale (BSS; Burnett, 1994); Myself-As-A-Learner (MALS, Burden, 1998); Taking Control of My Own Learning (Cameron, 1999); The Self-Regulated Learning Interview Schedule (SRLIS; Zimmerman & Martinez – Pons 1986; 1988 in Zimmerman, 1998) and The Inventory About Learning Approaches (Norwich, 1998). More specific learning and cognitive tools were also explored including the Learning Styles Inventory, (LSI; Kolb, 1976; 1985); the Learning Styles Questionnaire (LSQ; Honey & Mumford, 1982; 1986) and Cognitive

Styles Analysis (CSA; Riding, 1991). In addition to this, assessment measures from action-based research in thinking skills across the country were also considered. All measurements were reviewed in terms of reliability and validity statistics, the administration procedure and data generated.

Measures from current pieces of action research and devices such as Taking Control of My Own Learning (Cameron, 1999) were rejected on the basis that they had no comparative data nor reliability or validity statistics to standardise the instrument. Similarly criticisms surrounding the predictive validity of the learning styles questionnaire (Allinson & Hayes, 1998) and controversy over the internal consistency and construct validity from the Learning Styles Inventory (Allinson & Hayes, 1998; 1990; Wilson, 1986) meant they were omitted from selection. Other devices were considered inappropriate due to individual administration as in the Inventory about Learning Approaches and Self-Regulated Learning Interview Schedule, alongside the computer-administered Cognitive Styles Analysis. Although, all devices provided a rich source of data it was thought they would be too time consuming to implement given the sample size. Finally, the actual data yielded from measurements meant the Myself-As-A-Learner Scale was selected over the Burnett Self Scale. The former focused most closely on student's academic self concept as opposed to the latter which covered more general self concepts such as physical appearance, physical ability and peer relations. Furthermore the Myself-As-A-Learner Scale was an economical and effective measure to administer and score given the size of the sample. This combined with reliability and validity statistics to mean the properties of the tool were

suitable for large-scale research. The use of such a measure is the first attempt to examine student self-perceptions about thinking and learning, an area untouched in the research to date (Burnett & Proctor, 2002; DfEE, 1999).

The measurements selected aimed to ascertain the impact on learners across contexts, an area limited in previous research by an exclusive focus on attainment alone (Blagg, 1991; DfEE, 1999; Wilson, 2000). Likewise attention has been paid to validity, reliability and standardisation of instruments to provide a robust and scientific study as called for by De Corte, 2002; Gorodestsky et al., 2002; Higgins et al., 2002 and Watkins et al., 2001. Issues of transfer and maintenance have also been addressed which are areas untouched in present research (Blagg, 1991; McGuinness, 1999).

Sample

With regards to the sample, two key issues were considered in consultation with Dr. Reed at UCL. These were namely the sample chosen for study and the allocation of subjects to conditions. Prior research into thinking skills programs relied on small self-selected skewed samples of teachers under optimal learning conditions (McGuinness, 1999; TES, 2002; Wilson, 2000). It was considered essential in this study, therefore, that such recruitment bias would be minimised. This would be achieved by selecting the school and then year group for intervention rather than requesting individual teacher involvement. Teachers would be approached after schools had agreed to participate in the project. In a similar manner the piece will be original in

studying a varied population not restricted to one geographical location alone as in earlier research (Adey et al., 2002; Blagg, 1991; DfEE, 1999).

This research proposes to use a cross section of schools so as to include a variety of different settings and contexts, in contrast to previous work. Furthermore, demographic characteristics, an area also unspecified in prior study, will be used across a range of criteria including social economic status, number of children on the SEN register, and SAT scores to match schools effectively and ensure a representative sample. In this manner the criticisms levelled at research to date with regards limited sample size and scope can be addressed (Cotton, 1991; McGuinness, 1999; TES, 2002; Wilson, 2000).

Invitations for all schools in the region to participate in the project will be made. Selection will then be based on ensuring a representative sample including the variety and diversity of the national population on the criteria mentioned above. Subject sampling will occur to ensure a representative population. Systematic bias will be minimised by matching schools on criteria from the *intervention* and *waiting list control* group. Moreover the random allocation of schools in either of these groupings will avoid the impact of selection bias, whilst addressing ethical concerns of group allocation. The use of such procedures means the study is original in using more rigorous sampling techniques, a criticism made of previous studies (McGuinness, 1999; Wilson, 2000). Procedures for obtaining a representative sample and

random allocation to groups meant both aspects of the sampling design were addressed, unlike earlier work in the thinking skills domain (DfEE, 1999).

Procedure

Consultation with research teams in Northern Ireland, London and Scotland highlighted the issue of ensuring data collection was as child friendly, yet as standardised as possible. This assisted in practical terms in considering measurement administration. Firstly all studies had found the use of a script helpful in reassuring children participants as to the purpose of assessment whilst also reducing anxiety. As a result of this it was decided a script would be generated to highlight the purpose and process of the activities for children, whilst simultaneously reducing concerns over assessments. This would make procedures explicit to the participants and also ensure each researcher was presenting items in a uniform manner. The scripts would be extended to all measurements and scenarios, from a request for help to test administration, in order to minimise researcher impact or bias. The rate and tone of delivering directions would also be specified to avoid researcher influence.

The administration of measures were also considered and timings decided to spread assessment over one school day. Such a procedure was appropriate so as to ensure all students received similar measurement assessment at a certain time of day. This avoided confounding variables of time of day on performance and meant all participants were present for all measurements used. Schools felt this was the most efficient use of time in the curriculum.

Gaps between test phases were timetabled in with set activities to allow participants to remain concentrated throughout the session. This concept was supported and agreed by schools with additional breaks built in. Consultation with schools would also occur as to the behaviour management policy in place. Researchers, it was agreed, would follow the school policy on behaviour so as to provide consistency in rewards and sanctions given in the class setting. Schools and parents were aware of the procedures and processes to be used and agreed to them prior to administration.

Issues of transfer and maintenance have been ignored in earlier thinking skills research (Blagg, 1991; McGuinness, 1999; Wilson, 2000). Subsequently this study aims to build in procedures to address these issues. In particular the initial training of teachers provided through the ACTS intervention will be complemented by regular workshops in clusters of schools to support the use of the approach. These will use concepts such as Coaching, (Kearney, 1994) and Peer Group Supervision (Proctor & Inskipp, 1991) to support the development of the methodology. Likewise, clear links between school development and education development plans will be made to ensure inclusion and maintenance of ideals on a wider organisational level.

Issues and Implications for the Main Study

The preparatory phase of the research methodology aimed to address issues identified in earlier work in the domain of thinking skills study. In particular these factors related to rectifying the methodological limitations of prior studies, alongside examining unanswered practice questions. During the preparatory phase such areas were explored through the theoretical and practical aims, the chief implications of which follow:-

Methodological Issues and Implications

Current studies in the area of thinking skills have been limited by methodological and measurement flaws inherent in research design. These relate to the inadequate use of control groups, insufficient detail regarding measurement selection and restricted study scope (DfEE, 1999; McGuinness, 2003; TES, 2002; Wilson, 2000). The preparatory phase aimed to resolve these issues via theoretical consideration of the focus and design of the main study. With regards to research design a *waiting list control group* was agreed as a scientific method by which to evaluate effects using random allocation. This was ethically sound in that it delayed rather than denied an intervention, whilst controlling for a placebo effect of expectancy of improvement. Such a control group is original in thinking skills research as it moves away from a *no treatment control*, thereby, addressing earlier limitations (Blagg, 1991). Furthermore it provides an application of psychological experimental design to an area in need of more robust and scientific study (De Corte, 2002; Gorodetsky et al., 2002; Higgins et al., 2004;

Watkins et al., 2001). The focus of study was also sharpened to examine a three year period which addresses issues of transfer ignored in earlier work (Blagg, 1991; DfEE, 1999; McGuinness, 2004; Nisbet, 1993; Watkins et al., 2001).

In terms of measurement selection psychological understanding of the standardisation, reliability and validity of tools required to be scientifically rigorous assisted in decision making. Earlier work had been limited by devices chosen with little rationale and no reference to ethical considerations or technical detail. As Blagg (1991) highlights measurements to date *"have been either unreliable or inappropriate"* (p.26). The preparatory phase examined measurement devices in relation to ethical considerations to ascertain the sample on which standardised. Likewise it examined the validity and reliability to determine effectiveness and usage. Such an approach is rooted in psychology and is original in the area of thinking skills research. Moreover, the preparatory phase explored the measurement of student outcomes according to the psychological principle of triangulation based on the interactive factors framework (Morton & Frith, 1995). In this fashion the study is unique in examining multiple aspects of learner outcomes. Measurement selection, therefore, reviewed devices to tap in to cognitive, emotional and behavioural domains of participants' development.

Unresolved Issues and Implications

Earlier research has not adequately addressed the concept of transfer (Blagg, 1991; DfEE, 1999; McGuinness, 2003). The transfer of learning in

one context is critical in determining to what extent thinking skills taught, are then applied beyond the context of acquisition. Such an issue has been at the heart of the debate for some time (DfEE, 1999; Nisbet, 1993). This critical issue is to be addressed through the use of a three year study, tracking student development in a range of areas across contexts and time. Such issues are to be resolved in both the variety of measurements selected and procedures utilised. This focus on transfer and maintenance for both students and staff makes this research original in addressing a key unresolved issue (Blagg, 1991; McGuinness, 1999; McGuinness, 2004; Wilson, 2000). Previous research had not accommodated for such variables with *"difficulties of finding adequate ways of assessing the transfer of skills from the program to real-life context"* (Blagg, 1991) (p. 136). The preparatory phase identified the means to target this in both procedures and measurement selection.

Practical issues of research with regards to procedures were, likewise, explored in the preparatory phase. Consultation with schools and the sister project in Northern Ireland clarified administration of assessments with the development of scripts. Such devices incorporated ethical considerations to ensure the appropriate and most applicable implementation for child participants. Similarly the timings and prompt materials were designed to ensure an explicit process and rationale for students, staff and schools alike. A range of issues hereto undetected were also raised through consultations with UCL and other research projects. The issue of random allocation and subject selection were highlighted as two crucial components in sampling a

representative population. Additionally the need to minimise researcher bias in the speed of delivery and tone of assessment administration was identified.

Summary

The preparatory phase used the psychological research base to inform and develop a scientifically rigorous design which would be both replicable and ethically sound. In doing so it addressed limitations of previous research linked to inadequate controls and insufficient measurement selection. Moreover, the preparatory phase clarified the practicalities of the piece in terms of procedures and participants, with ethical issues considered. Finally the unresolved issues of transfer and maintenance were accommodated for alongside those hereto undetected. These factors combined meant the piece developed the size and scope of thinking skills research in a distinctive fashion.

Theoretical and practical considerations were explored in both the local and national contexts in this phase. The use of psychological literature in research design combined with consultations with UCL and with leading experts in the area informed the methodology and measurement selection. Likewise, it ensured the main study pursued future directions and resolved earlier limitations of research nationally. Consultation with schools, the LEA and other projects assisted in procedures being refined to best meet local need. Decision making from each of the key players meant the most relevant and rigorous measurements were adopted, alongside those of educational

applicability. Aims were also achieved within the constraints of time and finance without the need for a pilot. This was advantageous in ethical terms as no irrelevant assessment of children participants occurred.

Limitations of Preparatory Phase

Several limitations exist, however, in relation to the design, measurements and procedures of the preparatory phase which restricts its usefulness. In terms of design the preparatory phase was flawed, as in other thinking skills research, in that no actual pilot occurred. From a scientific standpoint this allows a number of confounding variables to be introduced into the design. Such factors include the unknown effect of researcher bias with inter-rater reliability unmeasured. Similarly the lack of rehearsal of the procedures of administration may confound the main study. The need for practice in procedures is an area which suffers a lack of objectivity with the experimenter effect unknown. In addition, the complexity of variables affecting measurement outcomes have not been examined such as the time of day, seating arrangements of participants and day of the week. An actual trial of the procedure would have investigated these factors whilst also gauging effectiveness in practical terms as to participants' engagement.

With regards to measurements in particular, none of the devices were trialled in practice. This may mean they are both inappropriate for participants and ineffective in ascertaining change. Such a lack of practice could mean timings of procedures are estimates alone. Issues also exist in terms of the extent to which measurements accurately tap into what they espouse. In

addition, as Topping (2003) identifies *"the main problem in measuring improvement in reading and thinking abilities is finding some measure which is accessible to young children, reliable and valid, sensitive to short term gains, and also economical in time for administration and scoring and in purchase cost"* (see website for detail). These factors and the limited tools available further reduced measurement selection. The impact of other variables interacting with these measurements is also unexplored from the social context (Meadows, 1998), to the experiences and beliefs of children (Busato et al., 1998). These measurement issues were not thoroughly examined in the preparatory phase. Finally, more specific measurement devices for certain thinking skills or self-esteem were left untouched, limiting the scope of study.

In future the preparatory phase would need to include some elements of a more traditional pilot study to ascertain practice implications for the main study. In particular a trial of the procedures and measurement devices could be utilised to gauge the instructional relevance and appropriateness for the study's hypothesis and participants. Schools could have been approached and given some incentive to engage in such a pilot. Such a design would have meant the researcher effect and inter-rater reliability could also have been estimated. In addition to this, visits to other similar projects may have been informative with regards measurement selection and administration procedures. Finally, with regards the limited measurements available, earlier work could have involved the design and standardisation of instruments to accurately tap into the key variables identified. These improvements could

have provided a more detailed evidence base for the main study. Such activities, however, were considered too time consuming for the present study given the restrictions of the standard fund bid.

Appendix 2; Technical detail of quantitative measurements.

Cognitive Abilities Tests (CAT3)

The CAT3 was standardised in the UK in autumn 2000 using schools selected by stratified random sampling. 331 schools from the primary age range participated and 225 from the secondary age range. At the same time an equating study occurred of 422 schools to equate CAT3 scores with CAT2E (second edition) from the 1984 standardisation. The validity of the measure is evidenced from two domains; both the inter-relationship between subtests and batteries and secondly the correlations to other forms of intellectual ability. With regard to the factorial structure of the nine subtests of the CAT3, correlations are at 0.7 with the lowest values for Verbal – Non Verbal and highest values for the Quantitative – Non Verbal correlation. Factor analysis used an oblique rotation using the Promax method. It is reported that the obliquity from the correlations between the three factors range from 0.6 to 0.7. In addition the first factor is said to account for around 64% of the variance and the three-factor solution accounts for 80% of the variance.

In terms of concurrent validity, correlation coefficients were taken alongside CAT3 Scores and Teacher Assessments of National Curriculum levels in the standardisation. Correlation coefficients were high, particularly in terms of mathematics teacher assessment with the Quantitative Battery at (0.50 – 0.71) and in English Teacher assessment and the Verbal Battery (0.53 – 0.67). The non verbal battery correlations were lower, perhaps due to the limited emphasis of this in the teacher assessment domains analysed.

Myself As A Learner Scale (MALS)

Standardisation occurred with sample of 389 (217 boys and 172 girls) in Year 7 and 8 in an urban comprehensive school. Norms were generated from this to allow for average range, mean scores and standard error and deviation to be ascertained. Validity of the measure was originally confirmed by generating items from informal interviews with students across the age range and then subjecting the scale to rotated and unrotated factor analysis. The rotated factor matrix yielded confirmatory positive results with the three strongest factors accounting for 43% of the variance and the remainder the other seven factors. The unrotated analysis indicated a strong first factor making 27% of variance with 18 correlations above 0.3. Five factors made 54% of the variance and 10 factors for 74%. Concurrent validity was ascertained by comparison with subscales of the Multidimensional Measure of Children's Perception of Control Scale (MMCPC) (Connell, 1985). Figures reported correlation at the 0.001 level; MALS with unknown cognitive control – 0.28, with control by powerful others – 0.16, and with internal cognitive control + 0.12. These significant results indicate concurrent validity. The measure has strong internal consistency with an alpha reliability index of 0.85.

The measure is unique in focusing on students self concept of themselves as learners and as such cannot be compared to other measures. It has been correlated with Cognitive Abilities Tests (CATs, Thorndike & Hagen, 1986) on the original standardisation sample. The results indicated a strong positive relationship between MALS scores at 0.001 level with verbal reasoning 0.41 and non-verbal reasoning 0.35.

Taxonomy of Problematic Social Situations (TOPS)

The Taxonomy of Problematic Social Situations originates on responses generated from 50 primary school teachers alongside 6 clinical psychologists to detect social situations in which students were frequently likely to experience peer relationship difficulties. Factor analysis then used the items to create six categories of social situations, which had high internal consistency reliability (Cronbuch's coefficient alpha). Test re-test reliability over a 6-month period was 0.79 for the overall TOPS score, 0.57 to 0.72 for category scores and 0.31 to 0.73 for item scores (Dodge et al., 1985). From this it was decided analysis would focus on category and overall scores alone, as opposed to individual items given concerns surrounding reliability.

The measure was originally used on a sample of 620 students from the ages 7 to 10 years in the USA. Two groups emerged namely the "adaptive" and "rejected" children from their peers. Findings illustrate differences between mean ratings between the "rejected" and "adaptive" pupils which confirm the construct validity of the measure. Moreover, standard deviations from the ratings of the rejected group were larger than those of the adaptive group across categories. An additional study by Nangle et al (1994) reported high temporal stability of the total and category scores over an 8-week period. This was conducted with 30 students aged between 10 and 11 in the USA. The measure, therefore, had reliability and validity statistics to make it a potentially useful device into examining students' social and behavioural outcomes, as opposed to the intrusive and less rigorous observational approach.

Appendix 3; Standardised script for administration of Cognitive Ability Scales (CATs).

Script for Assessment Session of ACTS

Good Morning, class ... my name is Jess Hudson / Janet Wallace and I'm an Educational Psychologist. This means I visit lots of children and teachers in schools to help with their learning. Today we are going to do some special work on how you think and learn. These aren't tests, but activities to explore how you learn. What we'd like you to do is to try your hardest, you might find some of the questions difficult, but just answer the ones you feel you can. If you have any questions do put your hand up and the teacher, LSA or myself will come over. We'll be breaking throughout the day for games and playtimes. Does anybody have any questions at this point?

Before we begin while I hand out the papers can everyone check they have a pencil and rubber on the table in front of them. When you have done this could you look at me and show me that you are ready to begin by being silent.

To begin with I want everyone to write their name down using a very special way.

Appendix 4; NfER Nelson administration procedures for Cognitive Ability Scales (CATs).

General Arrangements for Testing

Getting ready for the testing session:

1. Seat pupils so that they cannot easily copy from one another
2. Be sure the testing room is comfortable and well lit and that pupils have room to handle both the pupil's question book and the OMR answer sheet
3. Place a 'TESTING – DO NOT DISTURB' sign on the door
4. Try to forestall any interruption of the testing session by visitors or announcements
5. Ensure you have all the materials required
6. Before giving the instructions for completing the information panel on the answer sheet, write the following details in *capitals* on the blackboard – name of school, class and date of test (in numbers with a leading zero if needed for the day, the first three letters for the month and all four digits for the year)

And after the testing session has begun:

7. Move around the room while pupils are completing the name block and date of birth to determine whether they are doing it correctly. Pay particular attention to those pupils who usually have difficulty following instructions
8. Make sure pupils understand the example and practice questions before you start each test. Help any pupil having difficulty by repeating or rephrasing the explanations as necessary.
9. Move around among the pupils while they are taking each test to make sure that they are working on the correct page and filling in the answer sheets correctly.
10. Provide a short rest period between each test and the next

Appendix 5; Standardised script for administration of activities between subtests of the Cognitive Abilities Scales (CATs).

Activities to Use In Between Task Within Subtests

◆ **Verbal Subtest** Describe – Demo – Practice

1st task 10 mins

Rub a Dubs
Gently rub your hand in a circle on your tummy. Stop, then pat your head with your other hand gently. Now do both at the same time and at a similar pace. You should be rubbing your tummy whilst patting your head.

2nd task 10 mins

Concentric Circles

Children put their index fingers to touch at the top of an imaginary circle in front of the face. They circle their fingers in opposite directions simultaneously trying to keep perfect concentric circles. Their fingers should therefore meet at the top and bottom of the circle.

3rd task 10 mins

Game of Class Choosing

◆ **Non Verbal Subtest**

1st task 10 mins

Lazy Eights

With arm extended in front of you and your thumb pointing upwards trace the shape of an eight in the air. The eight should be on its side and as you trace it out in large slow movements focus your eyes on your thumb. Without moving your head trace three eights successively in larger movements. Then do it with your other hand and clap them together and do both.

2nd task 10 mins

Nose 'n' Ears

Children put their right hand across the front of the face and hold the left ear lightly. As they take the hand away they put their left hand across the front of the face to hold the right ear lightly. Then they swap! And again and again!

3rd task

Game of Choice For Class

◆ **Quantitative Subtest**

1st task 10 mins

Names in the Air

With your preferred hand write out your full name in the air. Use large movements. Do it forwards and backwards. Now use your other hand to write your name with both hands simultaneously (at the same time). If you are right-handed, start in the centre and work out. If you are left-handed start at the outside and work in.

2nd task 10 mins

Cross Crawl

Children lift the left knee and touch with the right hand, then right knee to the left hand and so on. With time they can then have a go at elbow to knees of opposite sides.

3rd task 10 mins

Game of Choice of Class

Appendix 6; Standardised script for administration of Myself as a Learner Scale (MALS).

MYSELF AS LEARNER SCALE (MALS)

NFER-NELSON

INFORMING YOUR DECISION

HOW I SEE MYSELF

Instructions: On the next page you will be given 20 questions to answer. Their purpose is to find out how you see yourself when it comes to learning and school work. Some people see themselves as being very good at learning and doing hard work, but others don't. We want to know what you think about yourself.

This is not a test. There are no right or wrong answers, so please try to answer the questions as truthfully as you can. Your answers will not be shown to anyone else.

First of all we need some information about you.

Name .....

Boy or girl.....

Date of birth .....

Today's date .....

Your age .....

Please read the statements carefully.

If you definitely agree, please put a circle around	a
If you agree a bit, but not so strongly, please put a circle around	b
If you think that the statement is true about half the time, please put a circle around	c
If you don't agree, please put a circle around	d
If you strongly disagree, please put a circle around	e

	Definitely agree	Agree a bit	True half the time	Don't agree	Strongly disagree
1. I'm good at doing tests.	a	b	c	d	e
2. I like having problems to solve.	a	b	c	d	e
3. When I'm given new work to do, I usually feel confident I can do it.	a	b	c	d	e
4. Thinking carefully about your work helps you to do it better.	a	b	c	d	e
5. I'm good at discussing things.	a	b	c	d	e
6. I need lots of help with my work.	a	b	c	d	e
7. I like having difficult work to do.	a	b	c	d	e
8. I get anxious when I have to do new work.	a	b	c	d	e
9. I think that problem-solving is fun.	a	b	c	d	e
10. When I get stuck with my work I can usually work out what to do next.	a	b	c	d	e
11. Learning is easy.	a	b	c	d	e
12. I'm not very good at solving problems.	a	b	c	d	e
13. I know the meaning of lots of words.	a	b	c	d	e
14. I usually think carefully about what I've got to do.	a	b	c	d	e
15. I know how to solve the problems that I meet.	a	b	c	d	e
16. I find a lot of schoolwork difficult.	a	b	c	d	e
17. I'm clever.	a	b	c	d	e
18. I know how to be a good learner.	a	b	c	d	e
19. I like using my brain.	a	b	c	d	e
20. Learning is difficult.	a	b	c	d	e

©Burden, 1998. Reproduced by kind permission of the author.
This measure is part of *Psychology in Education Portfolio* edited by Norah Frederickson and R.J. (Sean) Cameron. Once the invoice has been paid, it may be photocopied for use within the purchasing institution only. Published by The NFER-NELSON Publishing Company Ltd, Darville House, 2 Oxford Road East, Windsor, Berkshire SL4 1DF, UK
Code 009000 7300

Appendix 7; Directions for administration of the Taxonomy of Problematic Social Situation (TOPS).

TAXONOMY OF PROBLEMATIC SOCIAL SITUATIONS FOR CHILDREN (TOPS)

NFER·NELSON

Child's name: ..... D.O.B: .....

School: ..... Year: ..... Class: .....

Completed by: ..... Position: ..... Date: .....

Directions: This scale attempts to identify the kinds of situations that are most likely to cause problems for this child. For each situation, please rate how likely this child is to respond in an inappropriate manner (by hitting peers, aggressive verbally, crying, disrupting the group, withdrawing, appealing to the teacher for help, or behaving in some other immature, unacceptable, and unsuccessful way). In other words, how much of a problem is this situation for this child? This information can be used in designing the most effective intervention possible.

Use the following scale to answer:

Circle *1* if this situation is *never* a problem for this child.

Circle *2* if this situation is *rarely* a problem for this child.

Circle *3* if this situation is *sometimes* a problem for this child.

Circle *4* if this situation is *usually* a problem for this child.

Circle *5* if this situation is *almost always* a problem for this child.

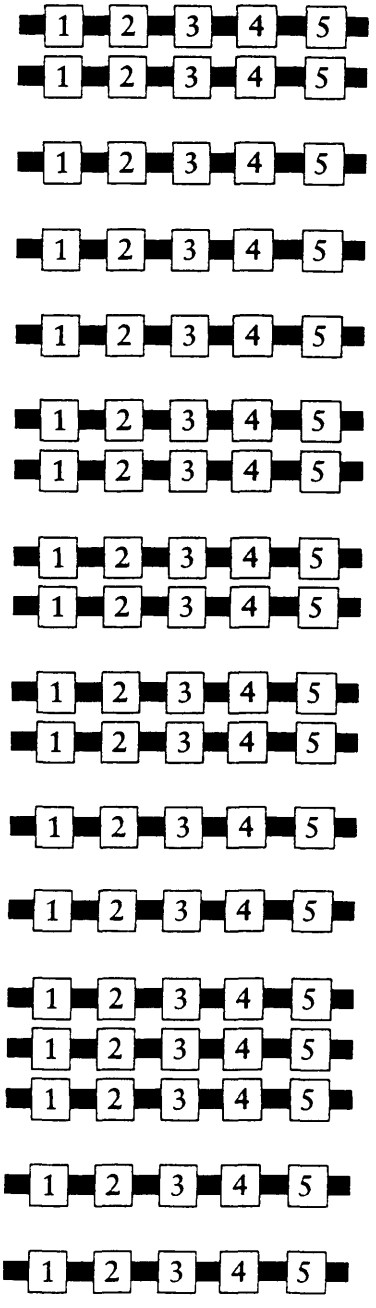
For example: Item 20: When this child is teased by peers. If you feel that, when this child is teased by peers, he or she almost always responds inappropriately or ineffectively (such as by crying), you would agree that this is a problem situation for this child and would circle 5. If you feel that when this situation occurs this child almost always responds in an effective and appropriate manner (such as ignoring the teasing), you would agree that this is not a problem situation for this child and would circle 1. *Remember, we are less interested in how frequently this situation occurs, and more interested in this child's response when it does occur.*

Does this child experience problems in these situations?

	Never	Rarely	Sometimes	Usually	Almost always
1. When this child is working on a class project that requires sharing or co-operation.	1	2	3	4	5
2. When peers notice that this child is somehow different (for example, wearing peculiar clothes, or walking strangely).	1	2	3	4	5
3. When this child has won a game against a peer.	1	2	3	4	5
4. When a peer takes this child's turn during a game.	1	2	3	4	5
5. When this child is playing a game with a peer and realises that the peer is about to win.	1	2	3	4	5
6. When peers call this child a bad name.	1	2	3	4	5

	Never	Rarely	Sometimes	Usually	Almost always
7. When a peer is allowed a privilege (such as winning a prize or standing first in line) that this child cannot enjoy.	1	2	3	4	5
8. When a peer performs better than this child in a game.	1	2	3	4	5
9. When this child asks a peer to play and the peer chooses to play with a third child instead.	1	2	3	4	5
10. When a peer performs better than this child at school work.	1	2	3	4	5
11. When peers laugh at this child for having difficulty in a game or play activity.	1	2	3	4	5
12. When this child performs better than a peer in a game.	1	2	3	4	5
13. When peers laugh at this child for having difficulty with a school work problem.	1	2	3	4	5
14. When this child performs better than a peer at school work.	1	2	3	4	5
15. When this child is having difficulty with a particular school work problem.	1	2	3	4	5
16. When a peer has something belonging to this child, and this child wants it back.	1	2	3	4	5
17. When this child finds out that he or she has been left out of a group, game or activity of peers.	1	2	3	4	5
18. When this child has something belonging to a peer and the peer wants it back before this child is finished with it.	1	2	3	4	5
19. When this child is playing with a peer and the peer accidentally breaks this child's toy.	1	2	3	4	5
20. When this child is teased by peers.	1	2	3	4	5
21. When a group of peers have started a club or a group and have not included this child.	1	2	3	4	5
22. When this child wants to play with a group of peers who are already playing a game.	1	2	3	4	5
23. When this child tries to join in with a group of peers who are playing a game, and they tell him or her to wait until they are ready.	1	2	3	4	5
24. When this child is accidentally provoked by a peer (such as a peer who accidentally bumps into this child in a line).	1	2	3	4	5
25. When this child is asked by a peer to share his or her toy or game (or pencil, or some other object).	1	2	3	4	5
26. When the teacher asks this child to work on a class assignment that will take a long time and will be difficult.	1	2	3	4	5

27. When the teacher is trying to speak to the entire class.
28. When this child is standing in line with peers and must wait a long time.
29. When this child is in the playground and a teacher is not nearby.
30. When this child is in the classroom with peers and the teacher must leave the room for a short period of time.
31. When this child is seated at lunch with a group of peers and a teacher is not nearby.
32. When a peer tries to start a conversation with this child.
33. When this child is sad, and a peer asks him or her how he or she is feeling.
34. When a peer has a toy, game or object that this child wants.
35. When this child has an extra toy and a peer asks him or her to share it.
36. When a peer expresses anger at this child.
37. When a peer has performed quite well at a task and is deserving of a compliment from this child.
38. When a peer is troubled, worried or upset and needs comfort from this child.
39. When a peer has been helpful to this child, and this child should thank him or her.
40. When a peer cuts into a line in front of this child.
41. When a peer tries to talk with this child.
42. When this child has accidentally hurt a peer and should apologise.
43. When this child needs help from a peer and should ask for help.
44. When this child loses a game with peers.



©Dodge, MClaskey and Feldman, 1985. Reproduced by kind permission of the authors..
This measure is part of *Psychology in Education Portfolio* edited by Norah Frederickson and R.J. (Seán) Cameron. Once the invoice has been paid, it may be photocopied for use **within the purchasing institution only**. Published by The NFER-NELSON Publishing Company Ltd, Darville House, 2 Oxford Road East, Windsor, Berkshire SL4 1DF, UK.
Code 009000 7302

TAXONOMY OF PROBLEMATIC SOCIAL SITUATIONS FOR CHILDREN (TOPS)

SCORING SHEET

NFER-NELSON
INFORMING YOUR DECISIONS

Child's name: ..... Date: .....

Directions: Enter the rating for each item in the box provided under the category to which it belongs. Add the ratings given to the items in each category of problematic social situations and divide the total by the number of items in the category to get the mean rating for each category. To calculate the mean rating for the whole taxonomy, add together all the category scores and divide by 44.

Peer group entry	Response to provocation	Response to failure	Response to success	Social expectations	Teacher expectations
				1.	
		2.			
			3.		
	4.				
		5.			
	6.				
		7.			
		8.			
9.					
		10.			
		11.			
			12.		
		13.			
			14.		
		15.			
	16.				
17.					
	18.				
	19.				
	20.				
21.					
22.					
23.					
	24.				
				25.	
					26.
					27.
					28.
					29.
					30.
					31.
				32.	
				33.	
	34.				
				35.	
	36.				
				37.	
				38.	
				39.	
	40.				
				41.	
				42.	
				43.	
		44.			

TOPS Total

Total _____ + _____ + _____ + _____ + _____ + _____ = ÷44 = ☐

÷5 = ☐ ÷10 = ☐ ÷9 = ☐ ÷3 = ☐ ÷11 = ☐ ÷6 = ☐

This scoring sheet is part of *Psychology in Education Portfolio* edited by Norah Frederickson and R.J. (Seán) Cameron. Once the invoice has been paid, it may be photocopied for use within the purchasing institution only. Published by The NFER-NELSON Publishing Company Ltd, Darville House, 2 Oxford Road East, Windsor, Berkshire SL4 1DF, UK.
Code 009000 7302

Appendix 8; Questionnaire for Pupils Involved in ACTS (post intervention).

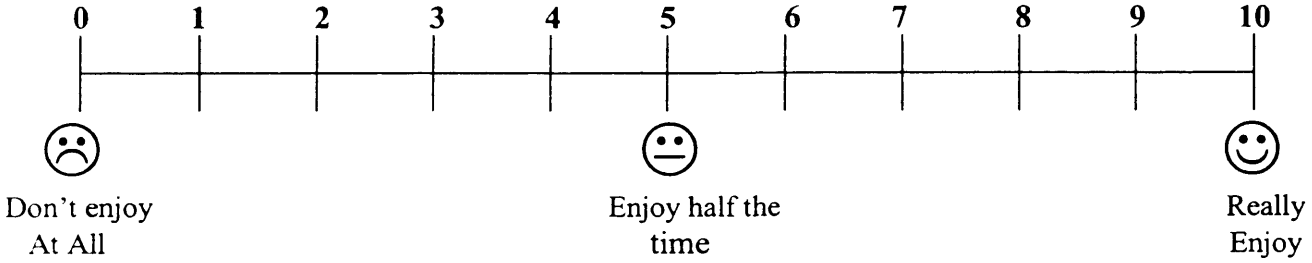
Questionnaire for Students Involved in ACTS

We're interested in your views about your thinking and learning. There aren't any "right" or "wrong" answers we'd just like to know about your opinions. Please fill out the questions below and ask an adult if you get stuck.

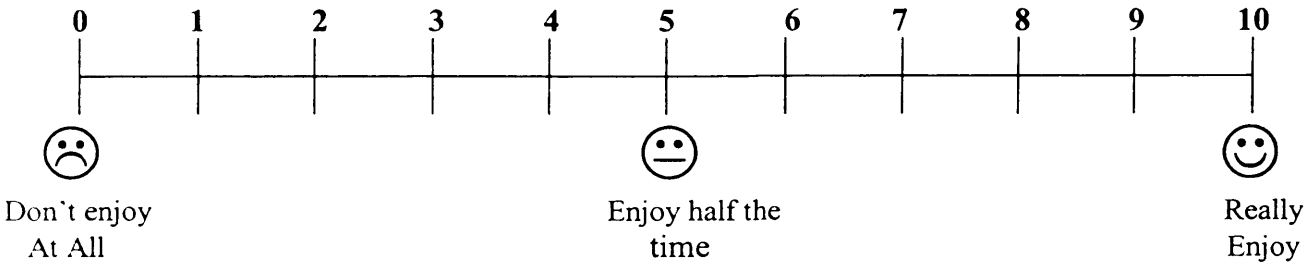
Name:-_____ **Class:-**_____

Name of School:-_____

1. How much do you enjoy everyday lessons in school (not thinking skills lessons)? Please circle.



2. How much do you enjoy thinking skill lessons? Please circle.

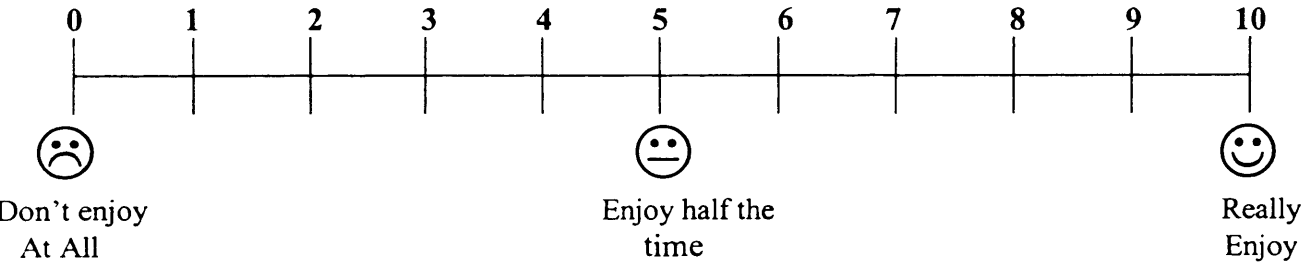


3. What thinking skills have you learnt this year?

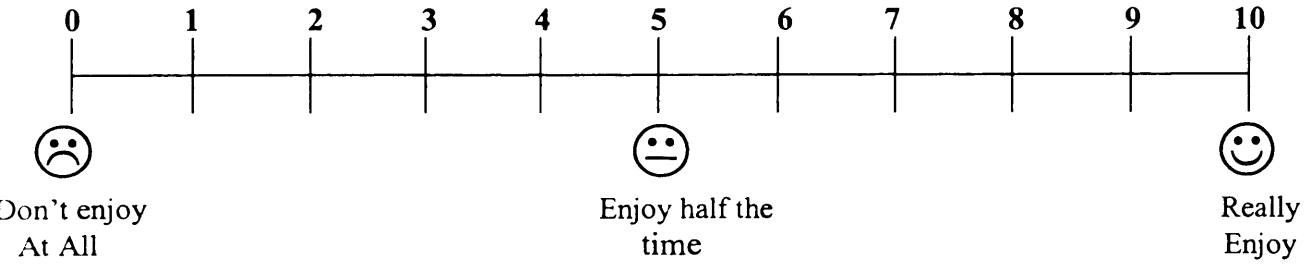
4. What are the good things about thinking skills lessons?

5. What are the bad things about thinking skills lessons?

6. How much do you enjoy working in your thinking group?



7. How much do you enjoy using thinking diagrams? Please circle



8. How have thinking skills helped you in other lessons?

9. How have thinking skills helped you outside school?

Thank you very much for your time.

Appendix 9; Questionnaire for Staff involved in ACTS (pre intervention training).

Experimental

QUESTIONNAIRE FOR TEACHERS INVOLVED IN ACTS

General Information

Date of Completion: _____ School: _____

Personal Data

Gender: _____ Religion: _____

Partnership Status:
Single [ ]
Married/living with partner [ ]
Divorced or separated [ ]
Widowed [ ]

Age: _____

Number of children living at home: _____

Professional Data

1. How long have you been in your current job? _____ Years _____ Months
2. How long have you taught year 4 children? _____ Years _____ Months
3. How long have you been a teacher? _____ Years _____ Months
4. When did you qualify as a teacher? _____
5. Have you received any thinking skills training or been involved in any related area? _____

   _____

   _____

   _____
6. What was your current knowledge base with regard to thinking skills prior to the training?

0 nothing at all	1	2	3	4	5 a great deal

7. What was your knowledge base with regard to thinking skills after training?

0 nothing at all	1	2	3	4	5 a great deal

8. Where would you like your knowledge base to be with regards thinking skills?

0 nothing at all	1	2	3	4	5 a great deal

Class Data

1. How many Year 4 children are in your class? _____

2. How many of these Year 4 children are at the following stages of the Code of Practice:

   School Action ____

   School Action Plus ____

   Statemented ____

3. How many of the Year 4 children have English as an Additional Language? ____

4. How many of these Year 4 children receive free school dinners? ____

5. Could you list the numbers of Year 4 children from each background:

   Asian ____

   White ____

   Black ____

   Other ____

6. How many of the Year 4 children have received any form of thinking skills instruction in the last two years? Please could you detail type of intervention, period of time involved and application to curriculum areas.

Thank you for your assistance in completing this.

Jessica Hudson
Educational Psychologist

Janet Wallace
Educational Psychologist

Appendix 10; Questionnaire for Teachers involved in ACTS (post training intervention – midshire version).

QUESTIONNAIRE FOR TEACHERS INVOLVED IN ACTS

Name	School
Gender	Number of Years Service

Please put tick in appropriate box:

1. What is your current knowledge base with regard to thinking skills?

0	1	2	3	4	5
Nothing at All					A Great deal

2. To what extent do you feel your knowledge base with regard to thinking skills has been extended through the ACTS training?

0	1	2	3	4	5
Not at All					A Great deal

3 What is your current skills base with regard to thinking skills?

0	1	2	3	4	5
Nothing at All					A Great deal

4. To what extent do you feel your skills base with regard to thinking skills has been extended through the ACTS training?

0	1	2	3	4	5
Not at All					A Great deal

5. How has ACTS contributed to your professional development as a teacher? (Please include positives and negatives).

Thank you very much for your assistance.

Janet Wallace
Educational Psychologist

Jessica Hudson
Educational Psychologist

Appendix 11; Questionnaire for Teachers involved in ACTS (post training intervention – sister project version).

ACTS in Surrey

End-of-Year Teacher Evaluation (July 2003)

We are now coming to the end of the year and you have been teaching thinking skills, designing infusion lessons and trying out different ideas in your classroom for almost nine months.

We would like to ask for your views on your experiences over the past few months that will help us to realistically evaluate what can be achieved with the ACTS methodology and to make improvements for the future. We are asking for your views in four main areas

Your views on the impact of ACTS on children's learning

Your views on the impact of ACTS on your own professional development

Your views on ACTS training days and training materials

Your future plans with regard to ACTS and teaching thinking generally

---

1. What do you think are the main benefits of teaching thinking skills through ACTs?
(Name at least 3 if you can)

2. What do you think are the main difficulties you encountered in teaching thinking skills? (Name at least 3 if you can)

3. Now that you have some experience of designing infusion lessons (finding contexts, etc), what advice would you give to a colleague who is just beginning...?

4. Explain how you used the thinking diagrams in your lessons.....

5. In what ways (if any) is your classroom different since you became involved in ACTS? (classroom management, group work, questions and questioning, atmosphere, children's level of engagement, etc)

6. What progress did you make in developing a vocabulary for talking about thinking, for evaluating and reflecting on children's learning and thinking in the classroom (metacognition)?

7. In what ways (if any) have your views of yourself as a teacher changed as a

result of being involved in ACTS?

8. Have you enjoyed teaching lessons that have a greater emphasis on thinking...? Explain.....

9. In your view, have the children enjoyed lessons that have a greater emphasis on thinking? Explain.....

ACTS Training Days and Training Materials

10. Could you comment on ACTS Training Days?

Strengths (things that we should definitely NOT change)

Improvements (things that could be improved)

11. Could you comment on the ACTS Handbook? How did you use it? How useful did you find it? What additional materials should be included?

12. Looking back, what additional support would you have found useful?

Future Plans

13. What plans (or even thoughts) do you have about incorporating ACTS into your lessons or schemes of work for next year? Explain...

14. Have you talked to any of your colleagues about ACTS? Explain.....

15. Does your school have any plans to incorporate ACTS on a wider scale (as far as you know...)?

16. Would you be willing to talk to other teachers more generally about the work you have been doing with ACTS (at teacher conferences, workshops, etc)?

17. For many different reasons we know that there have been constraints on your time this year and that there were differences in the degree to which you have been able to "run with the methodology" depending on your circumstances. In order to get some index of this, we are asking you to rate your level of engagement with the project (where level of engagement=number of lessons designed and taught, time spent planning, general involvement with the ideas).

Overall, how would you rate your **level of engagement** with the ideas and methods of the ACTS project ? Please be frank. Circle the number from 1-9 on the line below.

(Low) 1 2 3 4 5 6 7 8 9 (Very high)

Thank you for completing this evaluation.

We would be grateful if you could put your name on the form – for coding purposes. We will then make them anonymous.

Name_____

Appendix 12; Letter and Questionnaire for Head Teachers involved in ACTS (post training intervention).

Tel:
Fax:

[redacted]
Acting Head Teacher
[redacted] School
[redacted] Road
[redacted]

JH/js

30th June 2004

Dear Mr Bennett

Re: ACTS: Activating Children's Thinking Skills

We are writing to thank you for your schools participation in the ACTS project. The teacher's involved in the training have incorporated the methodology into their practice in an enthusiastic and thoughtful manner. They have also provided very positive feedback about the approach and its impact in their classrooms.

As part of our evaluation of the approach we are also interested in gaining your views on the effects of ACTS in your school. We would be very grateful if you would complete the attached questionnaire and return it to the above address by Friday 16th July.

Over the next year we hope to develop the methodology on a whole school approach. If this is an option that you as a school are considering, we would be happy to discuss this further with you.

Please do not hesitate to contact us if you have any further questions.

Yours sincerely

Janet Wallace
Educational Psychologist

Jessica Hudson
Educational Psychologist

Att.

INVESTOR IN PEOPLE Printed on recycled paper

QUESTIONNAIRE FOR HEADTEACHERS INVOLVED IN ACTS

We are interested in the impact of ACTS (Activating Children's Thinking Skills) in your school. We would be grateful if you could be as honest as possible to assist in the evaluation process. Responses may be used in the evaluation data, however, no school will be individually identified.

1. What impact do you feel the ACTS approach has had on individual pupils in Year 5 ?

2. Have there been any noticeable differences in classrooms where the ACTS approach is being taught ?

3. To what extent has the ACTS approach contributed to staff's professional development in terms of :

    a) Classroom Management and Structure

    b) Planning

    c) Teacher Effectiveness

4. Has there been any wider impact of the ACTS approach to the school as a whole ? If so, please specify.

5. How would you like to develop the ACTS approach in your school in the future?

6. Are there any other comments you would like to make on the approach?

Thank you for your assistance with this. Please return by Friday 16th July 2004.

Janet Wallace
Educational Psychologist

Jessica Hudson
Educational Psychologist

Appendix 13; Standardised parental letter to seek informed parental consent (pre and during intervention).

Tel:
Fax:

JH/kk September 2003

Dear Parents

We are undertaking the second year of a project in your child's school, which aims to develop children's thinking. We are exploring whether teaching thinking skills leads to enhanced attainment, reasoning and self-esteem.

Your child's teacher will be undergoing training in order to deliver the approach through the National Curriculum. This will not reduce the time spent on the National Curriculum but will provide teachers with an alternative way of teaching some lessons.

In order to evaluate the effectiveness of the approach, we would like to carry out some assessment work with all Year 5 children. This will take place in a whole class setting during a normal school day.

If you would prefer your child not to take part in these assessments or have any further queries please contact the class teacher as soon as possible. If we do not hear from you we will assume you give parental consent.

Thank you for your cooperation in this matter.

Yours faithfully,

JANET WALLACE
Educational Psychologist

JESSICA HUDSON
Educational Psychologist

INVESTOR IN PEOPLE

Professional Practice

Assignments

Jessica Lucy Dewey

D.Ed.Psy. (Volume 2)

2006

University College London

Contents

Professional Practice Assignment 1: Learning to Learn: An Evaluation of Thinking Skills Packages

Abstract

Thinking Skills is an area of considerable current interest in the realms of education, psychology, industry and politics alike (see DeCorte, 2002). This paper reviews thinking skills packages, which aim to develop children's thinking. Three major approaches are critiqued with theoretical orientations, characteristics and effectiveness ascertained. From this evaluation it appears that whilst successful approaches exist, substantial evaluation is still required to establish outcomes for students. The chief implications for current EP practice indicate a need for more scientifically rigorous research to provide an empirical evidence base for thinking skills programmes. Potential avenues of enquiry are highlighted alongside directions for future development.

Section 1: Aims and Scope of Assignment

The domain of thinking skills '*currently have a high profile*' (Topping, 2002). Indeed it has been claimed that '*the question of how to stimulate such higher-order intellectual processes in children has never been more widely discussed*' (TES, 14.6.02) (p. 24). The recent international conference on Thinking highlights the global importance of such a phenomena. Likewise, the recent explosion in thinking skills packages available indicates growing public interest (Adey, 2002; Gorodetsky et al., 2002). However as Gorodetsky et al., (2002) argue, schools are implementing such packages

with no attention to evaluation. This is cause for concern given the need for empirically supported interventions (Stoiber & Kratochwill, 2000). This paper therefore aims to provide a critical overview of the current thinking skills packages in schools. Such a phase in education was selected as it forms one of the core functions for educational psychologists (DfEE, July 2000). The purpose of this assignment is to examine the research basis and impact of such packages to ascertain effectiveness. Such evaluations will then be used to highlight implications for EP practice.

Definitions

As Wilson (2000) demonstrates, 'thinking skills' is an ambiguous term with little consensus as to its inherent skills. Furthermore Coles (1993) describes thinking as *'a vast and intricate family of activities'*. For the purposes of this assignment, the following definition will be used:

> '*The set of basic and advanced skills and sub-skills that govern a person's mental processes. These skills consist of knowledge, dispositions and metacognitive operations*' (Alvino, 1990) (p. 50).

This descriptor encompasses the idea that thinking includes 'higher level' processes such as critical thinking, decision-making and problem solving (Cotton, 1991). Moreover, it goes on to suggest that thinking is not merely knowledge acquisition but the capacity to reflect upon thinking and the processes involved (metacognition). Finally it incorporates the concept of

dispositions and affective aspects. Both metacognition and dispositions have been highlighted as crucial in developing thinking skills (McGuinness, 1999).

Overview

The focus of this assignment is to examine thinking skills packages and to suggest future possibilities for EP practice. Initially the piece will outline the context of such work and its relevance at a local, national, international and EP level. A critique of the existing research and literature will follow with psychological theory emphasised for three thinking skills packages. Finally the implications for further work in this area will be discussed with the role of EPs outlined.

Section 2 : Practice and Context

The rationale for investigating thinking skills packages exists at a number of levels. Initially local practice will be illustrated followed by a national and international focus. Finally current practice in the EP world will be ascertained to indicate the importance of continued development in this domain.

Local Context

In a large county Educational Psychology Service a thinking skills questionnaire was used to tap into the current knowledge, research and skills base of Educational psychologists. The anonymous questionnaire provided a sensitive and convenient measure of current practice. The questionnaire consisted of 10 items and used a range of Likert scales, closed option boxes

and open-ended responses. A copy of the questionnaire with an explanatory letter and return addressed envelope was sent to all Educational psychologists in June 2002 (see Appendix 1). A deadline date for return was stated providing respondents with a month. In total 29 EPs received questionnaires with 20 completing them, a response rate of 69%.

The survey findings illustrate a need in the local context to develop and expand the current knowledge and research base with regard to thinking skills. In relation to the question 'How much do you know about thinking skills?' 70% of EPs put themselves at or between 1 and 2 on a scale of 1 = nothing at all to 4 = a great deal. Likewise over 80% of EPs put themselves at a similar level in terms of their knowledge of the current research base. Qualitative comments supported this. With regard to the knowledge base concerning specific thinking skills packages over 95% were familiar with Instrumental Enrichment. However approaches such as subject specific and infusion were known only to 15%. This demonstrates a need to extend EP's knowledge in relation to the variety of thinking skills packages. Finally 90% of EPs who took part in the survey indicated they would like to implement thinking skills packages in schools.

At a local level the survey illustrates the knowledge and research base of EPs in the area of thinking skills and related programmes needs expanding. Furthermore there is an interest in implementing such programmes into schools highlighting a local need for this assignment. However there are several limitations with the survey, which restrict its usefulness. The chief of

these relate to the design and procedures used (See Appendix 2 for a detailed account).

Additionally at a local level there is standard fund money to develop and evaluate a thinking skills programme. A rationale for this assignment is to review current packages, which will form the basis of research in the LEA context.

National Context

On a national level the rationale for this assignment originates from four levels - society, schools, staff and students.

Society as a Whole

The importance of thinking skills at a societal level has increased with a need for knowledge construction capabilities (Gorodetsky et al., 2002). Such an interest has arisen in connection with the rapid change in technology and complexity of modern life (Cotton, 1991; Hatcher, 2002). As Wilson (2000) describes a broader range of competencies and 'higher order' thinking skills are necessitated given the rate at which information is expanding. This need for thinking skills to accommodate a rapidly changing world has likewise prompted governmental action (Pithers & Soden, 2000). In 1998 The Department of Education and Employment commissioned a report to overview thinking skills research. On the basis of this, in 2000 the Qualifications and Curriculum Authority incorporated five 'thinking skills' into the national curriculum. Most recently, 17 pilot authorities have trained key

stage 3 teachers in thinking which, is to be used across the country in Autumn 2002. There has also been an 'explosion of interest' in the teaching of thinking (Nisbet, 1993). As a result there exists a myriad of teaching packages to enhance students thinking, however, there is considerable debate about the merits of various approaches (McGuinness, 1999; TES, 2002; Topping, 2002). It is a purpose of this paper to critically evaluate such programmes to ascertain value.

Schools as Organisations

At a school level *'teaching children to become effective thinkers is increasingly recognised as an immediate goal of education'* (Cotton, 1991). Indeed there is growing consensus in the concept that schools should teach children how to learn as opposed to only 'imparting information' (Coles, 1993). Despite this attitude, however, research indicates that learning about learning is not a prominent feature of most classrooms (Hall et al., 1999 cited by Watkins et al., 2001). As Nisbet (1993) has argued, schools do use approaches to enhance learning in relation to specific areas e.g. exam preparation, however, they do not teach thinking as a transferable and generalisable skill. Moreover at a school effectiveness level, research demonstrates the positive impact to schools when a focus on learning has been adopted (Adey & Shayer, 1994; Gray et al., 1999 cited in Watkins at al, 2001; Munro, 1999). This paper aims to look at potential packages for thinking that could be used to meet such ends.

Staff as a Professional Group

There is growing research to suggest that the implementation of thinking skills packages can have a positive effect on teaching (McGuinness, 1999; Munro, 1999; Stewart & Smardon, 2002). Following the use of ACTS (Activating Children's Thinking Skills) teacher's felt that the programme had clarified and sharpened their concepts of thinking and had led to more effective planning (McGuinness, 1999). Moreover a study in New Zealand illustrated improved teaching and learning when a focus on thinking had been adopted (Stewart & Smardon, 2002). Finally, a study with secondary teachers indicated a programme focused on learning had a beneficial impact on the quality of teaching (Munro, 1999). There was an increase in the display of effective teacher behaviours, which were maintained over time. Such studies indicate the positive potential of thinking skills programmes on the quality of teaching. This assignment examines which programmes are associated with such promising effects.

Students as Individual Learners

In terms of the potential contribution of using thinking skills approaches with students, the research highlights a range of positive effects. Firstly, work by Baird (1986) indicated students had an increased awareness of the nature and purpose of learning. Similarly McGuinness et al's work (1997) shows how students were able to clarify thinking processes and became focused in their approach to thinking. Research suggests that, children's present knowledge in subjects improved but they were also more prepared for future content (Chi, 1996; Pramling, 1988). More objective measures demonstrate

the positive impact on academic performance from GCSE results to standardised measures of verbal and non-verbal reasoning (Adey & Shayer, 1994; Munro, 1999; Wegeif et al., 2002.) All this evidence reveals the potentially powerful impact of thinking skills approaches and this paper aims to examine such effects to determine the impact on students.

International Level

The impetus for thinking skills at an international level can be evidenced in events such as the 10th International Conference on Thinking (June 2002) which emphasised the need to develop thinking beyond current conceptualisations. As DeCorte (2002) argues, the educational gap is widening and learning is not developing at a rate commensurate with technology. In particular he highlights the need for more research to derive more reliable and generalisable conclusions. As Bond (2002) highlights there is a need for theory-driven research with a dialect between research and practice. Both these perspectives highlight the need for research internationally into thinking skills. This paper aims to provide a springboard for such research in a LEA context by previewing the current empirical basis.

Educational Psychologist Level

The study and research regarding thinking skills is of great professional interest to educational psychologists (Connor, 2002). Firstly there is the debate surrounding IQ and secondly there are the continuing controversies regarding stages of cognitive development. As McGuinness (1993) illustrates, the practice of teaching Thinking heralds *'new signs and*

developments in cognitive theory' (p. 305). Such issues are worthy of further exploration in relation to cognitive development theories and the nature of intelligence. It can be argued that educational psychologists are ideally placed to provide the relevant rigor to the study of thinking skills, being perfectly located to implement and evaluate thinking skills packages. As both Wilson (2000) and Topping (2002) indicate, evaluation studies at present are inconclusive. As McGuinness (1999) stresses *'considerable evaluation remains to be completed'*. There is also a need for empirically based practice with interventions supported by data (Bond, 2002; DeCorte, 2002; Fitzgibbon, 2002; Portwood, 2002). Both areas of conducting and designing research are core functions of educational psychologists (DfEE, 2000) and are worthy of future investigation (MacKay, 2002). This is an area ripe for development and the assignment intends to illustrate the crucial role EPs may play.

Section 3 : Psychological Theory and Research

The literature search tapped into Electronic-databases with search terms of 'thinking skills', 'thinking skills research' and 'children's thinking skills'. Manual searches of recent journals occurred alongside ancestral searches from overview articles. This was enhanced by attendance at the International Thinking Skills Conference at which experts presented current research.

The review aims to critique three major thinking skills approaches using Rudestam and Newton's (1992) critical evaluation checklist. The key programmes evaluated were selected as they each represent one of the

three models for developing thinking skills in schools (McGuinness, 1999; Topping, 2002; Wilson, 2000). Thinking skill approaches have been notably of three types, namely those which are subject specific, those that focus on general thinking skills, or those that are infused across the curriculum. The programmes were chosen over others as they have been highlighted as the most evaluated (Connor, 2002; Topping, 2002; Wilson, 2000). Likewise the research all took place with UK samples and as the DfEE review indicates, they are the most established programmes in the UK. Finally it was argued that an in depth critique of three key pieces of research would be more useful than a superficial analysis of dozens.

The review is divided into three sections. The first in a subject specific area (CASE - Cognitive Acceleration through Science Education), the second a structured general thinking programme (Instrumental Enrichment), and the third an infusion approach (ACTS - Activating Children's Thinking Skills). (See Appendix 3 Table 1 for overview). A final section highlights conclusions.

CASE (Cognitive Acceleration through Science Education)

1. Conceptualisation

Cognitive Acceleration through Science Education (CASE I: 1980-1983, CASE II: 1984-1987) had a clear rationale and issue to address. As Shayer and Adey (1993) point out CASE was a direct response to a study which, suggested that not more than 30% of adolescents develop early formal thinking by the age of 16. They highlighted that given secondary courses

require such thinking from at least year 9 (age 13-14), educators had a problem. The CASE projects were an attempt to improve cognitive development in order to increase the proportion of adolescents having formal operational capability by 14. The CASE project therefore had a clear problem identified, which had been illustrated in previous research (Shayer et al., 1976; Shayer & Wylam, 1978 cited in Shayer & Adey, 1993).

2. Theoretical Framework and Hypotheses

The CASE intervention had a detailed theoretical base drawing on a range of psychological theories. The approach was informed in part by Piaget and the notion of provoking cognitive conflict. CASE was also rooted in Feuerstein's work related to mediated learning experiences (MLE), Bruner's 'bridging' and Vygotsky's zone of proximal development. The only apparent limitation is that no justification is made as to why a subject specific approach was adopted. The criticisms of subject specific designs are that they do not allow for connections to be made across different areas of knowledge with issues of transfer and generalisation key. A rationale for such a subject specific approach would have been useful in addressing concerns. In the actual construction of hypotheses the research literature is scant. Hypotheses are not clearly testable, although, general expectations and relationships are evident.

3. Research Design

The research approach is clearly detailed with a rationale. As Shayer and Adey (1993) indicate the *'methodology was explicitly interventionist'*. The time frame of two years was justified by earlier research (Feuerstein et al., 1980). Research involved 12 classes with matched controls from two middle schools and one comprehensive. Students in the experimental group received one period (70 minutes) of Thinking Science every two weeks with the control group receiving ordinary science. This methodology had no rationale provided which prompts a number of criticisms. Firstly, although a control is used it essentially acts as a 'no treatment control.' It therefore suffers from the problem that the intervention group may have had the expectation of improvement and created a placebo effect. Similarly given there was an amount of time specified, an attention placebo control could have been effectively utilised. As it stands the research design does not adequately control for extraneous variables.

In terms of student variables, changes in operational thinking, science achievement tasks and GCSE were used. These are justified in relation to theoretical underpinnings. All of these variables relate to outcomes and performance on attainment tasks offering a very limited view of impact on learners. Likewise there is a lack of analysis in terms of the teacher, curriculum and lesson delivery. The fact that these variables are not included reduces the validity of conclusions drawn.

With regards to the sample, three schools were selected which indicates a large sample particularly given the long-term nature. However, no demographic characteristics are provided for schools or students alike. Both these factors mean the extent to which these results are generally applicable to the national population remains ambiguous. Also no justification is provided for why certain schools or teachers were selected. This once again limits the impact of the research for if schools self selected then this may have skewed results.

Of the measures used, no detail is given in terms of reliability of validity. The impact on students and staff attitudes and behaviours is not measured, thereby introducing a multitude of confounding variables. No pilot was used in trialling the measures, moreover some of the measures were used only following the intervention. Such factors confound research design considerably.

4. Results and Discussion

With regards to statistical results there was little rationale given for why a particular test was used. Furthermore the appropriateness of tests and confidence limits were unspecified. In handling the data several variables were unaccounted for. These are ignored in analysis and interpretation.

In drawing conclusions Adey and Shayer (1993, 1994) report striking results, yet immediate test results showed only a significant gain on reasoning tests for some boys. After a year the effects in science achievement test were statistically significant for some groups but were not stable across years or

gender. The most widely discussed results were those in terms of GCSE grades where gains of one grade were made. Shayer and Adey (1993) describe how this indicates that even after the CASE intervention the effects increased and 'can be claimed to be permanent'. Such a conclusion is not seen in the data, as some students did not show such gains. The results do not necessarily indicate permanent change or highlight CASE as the prime mechanism for such impact. Alternative conclusions that would be consistent with the data are not discussed. Few comments are made as to limitations, however, practical implications are discussed alongside further research and future directions. No reference is made of the general criticisms that subject specific approaches restrict thinking. As Coles (1993) argues the approach '*does not enable children to forge links between different areas of knowledge*' (p. 339). As such it represents a significant omission.

5. Summary

The CASE approach offers a context-dependent intervention with strong theoretical underpinnings. Likewise its longer term perspective and connections to previous research enhance its value to the empirical field. A drawback of the approach is the assumptions it makes concerning research design, which introduce a range of extraneous variables. Moreover, the gains and impact made are discussed only in relation to the success of CASE as opposed to any alternative explanations. Despite these shortcomings, however, the CASE approach is claimed to be one of the most carefully evaluated thinking packages (Topping, 2002).

FIE (Feuerstein's Instrumental Enrichment)

1. Conceptualisation

The major issue being investigated was an evaluation of Feuerstein's Instrumental Enrichment (FIE) programme by Blagg (1991). The evaluation aimed to assess cognitive and affective changes in pupils and their teachers. The rationale for investigation was explicit and made in connection with a range of contexts. Additionally limitations of previous research were highlighted and how this piece sought to resolve them. The concepts of instrumental enrichment from a theoretical, empirical, and practical perspective were highlighted (Blagg, 1991). Theories, assessment models and intervention programs were clearly explained with the rationale for exploration justified.

2. Theoretical Framework and Hypothesis

Broad aims and objectives were clearly made in evaluating the potential effects of FIE on the pupils. The main goals of FIE were analysed and translated into testable hypotheses. Blagg (1991) generated specific hypothesis, which divided Feuerstein's ideas into directly observable and testable constructs. Such hypotheses were clearly stated and directly linked to measures making them replicable. The only drawback was that the hypotheses were selected on the basis of measurements available rather than the theoretical basis. Hypotheses relating to students were all linked into Feuerstein's original hypotheses. However, hypotheses relating to staff members were not related directly to Feuerstein with relationships unclear.

3. Research Design

Blagg's study (1991) was a two-year intervention evaluation project. It was conducted over 2 years with 14-year-old low-achieving adolescents in 4 schools in Somerset. A clear framework to monitor the effects of Instrumental Enrichment was used. Control groups and experimental groups were selected within schools to reduce the impact of confounding variables linked to school characteristics. The only issue is the extent to which the control and experimental groups were matched. Instead it appears they were selected and nominated by teachers. This introduces a confounding variable into the equation, that of teacher perception and adds ambiguity to pupil selection.

The variables in relation to students were clearly delineated in accordance with Feuerstein's original hypothesis. In relation to the teachers as variables several areas were focused on, including personality, attitudes and teaching style. These aspects, although relevant, were not sufficiently justified.

The sample was located in a small industrial town and included all 4 secondary schools. The key difficulty is how representative that location is to the country. Such an issue could, have been minimised, if demographic characteristics had been ascertained and related to the national population. Similarly using a range of secondary schools across Somerset would obtain a more representative sample. In selecting the sample 'low-achieving' students were chosen on a government-based initiative. This is an ambiguous descriptor including a whole range of assumptions. Teachers

selected students on the basis of this, which may have significantly skewed sampling. In future student samples would need to be more objectively selected and appropriately matched between the experimental and control. Likewise teachers were self-selected which may have impacted on results.

Measures used to assess students were selected according to Feuerstein's original hypothesis. The rationale for measures used, however, was lacking. Moreover no reference was made to validity or reliability. Furthermore no pilot occurred and procedures of when measures were implemented were not detailed. In addition attitude and behavioural measures used observation schedules in which teachers rated students. Again no data on reliability or validity were given and the impact of teacher bias ignored. The teacher measures included a range of scales which '*were not factor analysed or checked for reliability and validity* ' (Blagg, 1991) (p. 45). This makes both use and value limited. The reason these were selected was also not addressed.

4. Results and Discussion

Basic analytic techniques were justified in relation to the data collected, however, confidence limits were unspecified. The other statistical issue is that students had several measures of ability and attainment. Given the vast number of these they may have reduced the importance of certain variables. Data analysis is consistent with what the data indicated and the way in which tests were set out meant interpretation was possible.

The key conclusions drawn reported no significant improvements in the FIE trained group compared to the control groups. Other potential conclusions, which fitted the data, were also discussed. With regards to changes in teachers some significant results did emerge in terms of developing positive attitudes about 'low-achieving' students and themselves as teachers. This conclusion was consistent with the data.

The limitations of the study were referred to including generalisation and context. As Blagg (1991) himself remarked '*although Instrumental Enrichment was probably not evaluated under optimal conditions, the context was certainly very favourable*' (p. 125). It is more likely that this research illustrates how successful such work would be in ordinary schools. One limitation, which was not highlighted sufficiently, was the disparity between schools in terms of control groups and support for FIE. The practical and research issues of the project are discussed with particular reference to the many variables operating in schools. The practicalities of both implementation and program design were also illustrated.

5 Summary

The UK Evaluation of Feuerstein's Instrumental Enrichment program provided a detailed and critical piece of research. It addressed many of the research issues highlighted at the time by providing an independent study with control groups monitoring the impact of an intervention over a significant period. Conceptualisations and hypothesis were clear with an experimental design to match. The key limitations related to measures utilised and

variable categories. Results and conclusions were clear indicating that FIE had no impact on students. Research and practical implications were appropriately discussed offering future perspectives. Despite some limitations, the study offered a realistic view of the use of a cognitive program in ordinary schools.

ACTS (Activating Children's Thinking Skills)

1. Conceptualisation

The primary purpose of ACTS was two fold; to develop the methodology and to gauge initial reactions of students and staff to the programme. The piece didn't aim to investigate any major problem or issue outlined in earlier research. Additionally the goals were general and vague in relation to intended outcomes or gains. The area of research was justified, because a rationale was provided for the benefits of infusion and the appropriateness of the approach for classrooms. Likewise the concept of thinking skills and infusion were rooted in earlier work (Swartz & Parks, 1994). ACTS also systemised a range of teaching strategies, which originated in prior research on powerful learning environments (De Corte, 1990). Such strategies had a clear rationale and empirical basis.

2. Theoretical Framework and Hypotheses

ACTS was an infusion approach that aimed to develop thinking skills across the curriculum. This ensures thinking is embedded into a range of meaningful contexts. The rationale for using infusion were clear and included

how transfer could be reinforced and how thinking skills could be matched to curriculum topics. A rationale for selection was given in relation to other approaches.

A research question was not evident but rather a general aim '*to promote the development of thinking skills in ordinary classrooms*' (McGuiness, 1999) (p.19). The purpose was to develop the methodology and to gauge initial reactions. As a result, there was no clearly stated research question with a rationale, the terms used were non-specific and immeasurable and there were no hypotheses present (making replications difficult).

3. Research Design

The research design was exploratory to gauge initial reactions to ACTS. The design was not clearly developed with a rationale. ACTS involved work with 17 KS2 teachers. The teachers attended 6 training days over a 6-month period. At the end of the year teachers completed an 18 item open-ended questionnaire about ACTS. As a result there was no clear research design with no evidence of any pre or post measures of students or staff. No controls for potentially confounding variables such as age or years of teacher experience etc were used. The lack of control group meant change could not be attributed to ACTS alone. The identified variables were student and staff reactions, however, these were not clarified. Similarly how reactions were to be measured was not clear.

The sample consisted of 17 upper primary teachers with no other reference made toward teacher or school characteristics. This meant potentially confounding variables were uncontrolled. The difficulty with this lack of information means it is unlikely that we can extrapolate from results found as we are unsure to what extent this sample is representative. Another important factor is that it is not reported why teachers were selected, this may impact on results.

The only measure mentioned was an 18 item open-ended questionnaire. There are several difficulties with both the use and basis of such an instrument. Firstly there are issues with the validity and reliability but also the fact no pilot occurred. The questionnaire was used following the programme and a year later. Therefore no pre measure was used to ascertain change. Also the questionnaire was administered at the end of training days, which as McGuinness et al., (1997) puts it was '*when spirits were high*.' This may have meant the measure captured opinion at an unrepresentative snapshot. Finally the study's purpose was to measure student and staff reactions, however, this measure only focused on staff perceptions. The impact on students was ascertained by questioning the staff, which introduces bias. Another measure was required to ascertain student change objectively.

4. Results and Discussion

No detail was provided for how questionnaires were analysed, rather it is reported that teachers evaluated the impact of ACTS on the students and themselves as positive. As it stands this information in qualitative form is

useful as an initial view into teacher reactions. It does not, however, provide any quantitative evidence as to the effect of ACTS and whether any changes in staff or students occurred. The data used for results is not appropriate to the design, as they do not refer to students. Furthermore no triangulation for any of the points made is given. For example observations and video evidence before, during and after the study could have supported teacher's perceptions. Work with students could also have investigated change. To conclude in the results section the potentially confounding variables of student, staff and school characteristics are not addressed.

The conclusions drawn from the study were that '*teachers were very satisfied overall with the ACTS Methodology* '(McGuinness, 2000) (p. 10). This was not, directly stated in the data. Instead the results indicated the perceived positive impact on staff. Alternative conclusions for the data were not discussed and accounted for. The limitations of the study are, however, discussed. In addition future directions were highlighted in terms of the need for more evaluation. The theoretical and practical implications were pinpointed.

5. Summary

ACTS provided an initial exploration into whether an infusion approach could be implemented and how teachers reacted to it. The original purposes of promoting thinking skills or gauging student reactions were not achieved. Although a strong theoretical framework was identified the lack of clear hypotheses limited what the study achieved. The research design had no

clear structure with variables ambiguous and potentially confounding. Similarly, the only measure used was not reported as valid or reliable. As a result the scope for the results was limited reporting on teacher perceptions alone. Despite these difficulties however, the work did succeed in developing the methodology in ordinary classrooms and provided insight into staff perceptions. Moreover, it highlighted theoretical and practical implications with future directions.

Summary of Psychological Literature

Current psychological research suggests that thinking skills packages may have some positive impact on both students and staff. There are, however, several issues with the way in which these domains have been explored, which require improvement.

The limitations of the literature reviewed pertain to issues of research design. Firstly, there were a range of confounding variables inherent in all studies meaning a number of factors were at work which were not adequately controlled for. Likewise the size and scope of the work was limited by small sample sizes and often optimum classroom contexts. As McGuinness (1999) points out '*problems with scaling up and transferring the effects to everyday classrooms have been identified*' (p. 29). Finally a range of measurement issues significantly impact on the validity of the research with measures selected with no rationale and no reference to reliability.

The main controversies that remain within the literature relate to the issue of transfer and practical implications. The transfer of learning is most critical in determining to what extent the thinking skills taught are then applied beyond the context of acquisition. Such an issue has been at the heart of the debate for some time (DfEE, 1999; Nisbet, 1993). Such factors are crucial but not adequately discussed in the literature. Other unresolved issues are the practicalities of the approaches in context (TES, 2002; Topping, 2002). Furthermore, these relate to larger issues of the nature of pedagogy and the curriculum which are untouched in the research.

Several unanswered questions are obvious in the present psychological literature. Firstly there is the issue of age and whether there is an appropriate, optimum point at which to introduce thinking skills. Then, there is the actual time of the thinking skills package. Additionally there is huge controversy over which type of programme should be used with the possibility of a mixed model an avenue still to be explored.

To conclude, the research literature, although indicative of positive effects is riddled with a range of methodological and measurement flaws which limit its validity. As Wilson (2000) highlights, '*evaluation studies are inconclusive*' with a very real need for empirical evidence. Research must occur in everyday classroom contexts with a clear research design and issues of transfer addressed. A range of psychologists support this perspective calling for more robust evidence (Gorodetsky et al., 2002; Higgins, 2002).

Section 4 : Integration of Theory, Research and Practice

It is evident from the context and the review of the psychological literature that a gulf exists between current practice and what research advocates. This section explores the way in which practice and research overlap and conflict. In the local context the need to increase EP's awareness is examined and recommendations made for short, medium and long-term action. The national perspective is then focused upon with the mismatch between current educational practice and research highlighted. The implications for Educational Psychologists are delineated in terms of empirical research to minimise previous limitations, answer important practice questions and pursue new directions.

Local Context

In the context of the local Educational Psychology Service, there exists some overlap between research and practice issues along several dimensions. Firstly as the local survey indicated there was a great deal of interest in implementing thinking skills packages in schools. Such a display of enthusiasm parallels the explosion of recent research in this field. Both context and literature indicate growing curiosity in the domain of thinking skills and highlight the need for more research in this sphere. Finally there is a general accumulated body of knowledge at both a practice and research level. There is an understanding of the existence of different kinds of thinking skills programmes which, aim to develop children's thinking using varying methodologies.

Large discrepancies exist between local practice and research issues along a range of criteria. One relates to the limited knowledge of thinking skills and research within the EPS, which contrasts markedly with the wealth of literature and expertise available in the research. As the survey indicated 70% of EPs put themselves at or between a 1 or 2 on a scale of 1 = nothing at all to 4 = a great deal, in relation to the question 'how much do you know about thinking skills?' This is in direct contrast to the detailed research base highlighted in the literature. It appears there is a shortfall between EPs knowledge in practice and the detailed information available in research. Such an imbalance needs re-addressing so practice is in line with current research.

Another area of conflict between practice and research pertains to the specifics of the various different thinking skills approaches. The research literature evidences the three main types of packages available. In practice only 15% of EPs knew about subject specific or infusion methodologies. This demonstrates a need to extend EPs knowledge and practice to a comparable level with the research so the variety of thinking skills packages are clear. Such information would be advantageous to Educational Psychologists in evaluating the types of approaches and impact. This would enable EPs to consult with schools as to different methodologies. Knowledge of the research findings may also assist EPs in using appropriate thinking programmes according to the strength of findings as opposed to how they are marketed. Likewise, specific knowledge would be useful in determining the best approach for varying purposes.

The great disparity between the practice context and research findings would need to be addressed. This could occur on several time-scales as outlined below:

- **Short Term Action – Increasing Awareness**

The increasing interest in thinking skills from society highlights the topic as an area of continuing professional development for the Educational Psychology Service. A special interest group could be set up with the aim of raising awareness and knowledge. Key information regarding the various methodologies and empirical evidence base could be collected by area teams and collated into a critique of the approaches alongside a basic comparison. In conjunction with this, guest speakers from each of the thinking skills approaches could be used. In this respect the purpose and effectiveness of such projects would be transparent to the service.

- **Medium Term Action – Develop Active Understanding**

The next phase would aim to extend EP's knowledge into workable practice by transferring the conceptual knowledge into everyday usage. A special interest group could explore the measures, materials and approaches highlighted in the research to ascertain application. Piloting such resources in everyday practice would be a measure of usefulness. A focus on the research could also generate the design of small-scale projects in ordinary classrooms. Action research with schools could occur, exploring the value of measures or specific techniques e.g. cognitive coaching strategies in areas such as spelling. This would provide a starting point from which to build a

more experienced and informed perspective on thinking skills. Such knowledge could be shared with the staff, school and service to further develop a dynamic understanding of thinking skills in real life contexts.

- **Long Term Action – Forward Knowledge Base**

Once the knowledge base had been extended and translated in small-scale practice larger projects could occur. One focus could be to develop and design assessment techniques with reliability and validity statistics for thinking skills and metacognition. Both research and practice indicate a need for objective measurement devices in this sphere (Blagg, 1991; Bond, 2002; Fisher, 2002).

Alternatively, large-scale project work could link up with key professionals to ascertain thinking skills packages effectiveness. Evaluation studies of each of these approaches could occur to ascertain impact on learners. Similarly links could be made to other Educational Psychologists engaged in such evaluative research. These would provide professional connections in developing understanding in the area of thinking skills research. Such work would aim to extend the empirical evidence base for cognitive interventions.

National Context

At the national level there is a clear overlap in the area of research and practice as to the importance and relevance of teaching thinking skills. In practice both schools and governments are keen to pursue the instruction of thinking (Gorodetsky et al., 2002; TES, 2002). This is paralleled in research,

which, proposes teaching thinking has positive affects. Furthermore the explosion of enthusiasm in practice contexts matches the recent increase in research (McGuinness, 1999; Nisbet, 1993). Despite this mirroring of interest there is conflict between research and practice issues that need addressing.

Current practice in the education system seems contrary to what would be advocated on the basis of research findings. It appears that National practice assumes that thinking skills packages are effective and are thus implemented. This contrasts markedly with research that indicates such effectiveness is not necessarily the case with evaluations remaining inconclusive (Wilson, 2000). As Coles (1993) points out, even in the research there is *'a good deal of scepticism about the notion that you can teach thinking at all'* (p. 341). There is considerable debate in research as to the impact and merits of various approaches (Topping, 2002). Moreover all research literature advocates the need for more empirical work with key effective thinking skills pinpointed (McGuinness, 1999; Watkins et al., 2001). Indeed, theory driven research is required with links between practice and research transparent (Bond, 2002; DeCorte, 2002). From the literature reviewed, the number of limitations and unanswered questions highlight the need for more research. Educational Psychologists are uniquely placed to engage in such research being on the interface of theory and practice with the requisite research skills. The following sections illustrate the role of EPs in engaging in scientifically rigorous research to reduce the shortfall between practice and literature.

◆ Addressing the Limitations of Previous Research

A review of the psychological literature highlighted a number of confounding variables inherent in research designs. The majority of studies used inadequate controls, which received no alternative treatment or time in comparison to the experimental group. Additionally insufficient detail meant replication was impossible. EPs could engage in research of a more rigorous nature to resolve such criticisms. Clearly delineated variables selected on the basis of theory and not measurement devices could be used. Likewise control groups could apply a waiting list, attention-placebo control or alternative intervention. Samples would need to be matched within the experimental and control groups and randomly allocated. Furthermore schools and students would be selected to form a representative sample of the national population.

The size and scope of the research would also need to be improved from the current base. As both Wilson (2000) and McGuinness (1999) highlighted most research has occurred in optimal learning environments. Research could focus on the impact in ordinary classrooms with randomly selected staff and schools. Similarly the range of children populations needs extending to include a range of needs. Topping (2002) points to the restricted focus of current work; '*virtually all the research literature is concerned with teacher-directed instruction*' (p. 287). Research could be extended by EPs to cover a range of relevant contexts.

Finally, a review of the issue of measurement devices in literature leaves much to be desired. Outcome measures were not linked to hypothesis or selected with a clear rationale. In future research EPs need to justify the selection of measures on the basis of reliability and validity data. Additionally links with both the hypothesis and theory base need clarifying. Pre and post testing with the same sample also needs to be delineated.

◆ Examining Practice Questions Unanswered in Previous Research

A key issue left relatively untouched in the research literature was the issue of transfer and the extent to which thinking learnt in a specific context transferred to other areas. Such a point has dogged the thinking skills research for some time (Blagg, 1991; McGuinness, 1999; Nisbet, 1993). EPs involved in constructing new research would need to assess learner outcomes across a range of contexts. The concepts of 'high road' and 'low road' transfer could be investigated (Salomon & Perkins, 1989). The durability of training over the short and long term could also occur with several post-test phases.

Another issue unanswered in the literature reviewed were the practicalities of such projects in context. Issues relating to changes in classrooms in terms of curricular design, materials and pedagogy are all relevant but as such un-investigated (McGuinness, 1999). Research would need to look more closely at the projects and liaise with staff, students and schools to ascertain wider impact. This could be achieved through more qualitative techniques such as

questionnaires, structured interviews and focus groups. Such information may yield important practice implications.

Finally there are several unexplored issues which, may determine the effectiveness of a thinking skills programme. The issue of optimum age to begin teaching thinking has been unanswered (Nisbet, 1993). Similarly the amount of time devoted to such activities and whether this has an impact is uncertain (Cotton 1991). EPs could engage in controlled studies to evaluate the potential impact of these variables. There is also the question surrounding infusion or enrichment programmes and which is more effective. Indeed in the research, no study attempted to compare the effectiveness of one programme with another. Evaluation studies could utilise 3 experimental groups of each type of programme with tightly matched samples in a mixed model design (Coles, 1993).

- **Future Research Possibilities Arising From Research and Practice Issues**

There exists a range of research possibilities stemming from the omissions in the literature review and practice issues raised. Firstly, there is scope for focused work in evaluating the role of emotions and dispositions of children to thinking. As McGuinness (1993) indicates '*the influences of affect, motivation and appraisal on learning and thinking are still largely unexplored*' (p. 309). Likewise student's perceptions warrant further analysis (McGuinness & Nisbet, 1991). The role of metacognition similarly could be concentrated upon. The social situation could be focused on with the role of

peer mediators or tutoring examined (Topping, 2002). Other areas to be investigated include the use of information and communication technologies (McGuinness, 1999; Wilson, 2000). Finally the development of thinking classrooms and schools are also an area of exploration (McGuinness, 2002).

Concluding Comments

This assignment illustrated the local, national and international interest in Thinking Skills. The literature review mirrored such a sentiment with numerous studies exploring how to teach thinking. Despite the existence of some successful elements, substantial evaluation is needed to ascertain the impact on students. The scope for minimising the current shortfalls between research and practice falls at a local and national EP level. It is crucial that Educational Psychologists become more knowledgeable and active in constructing research in this domain. EP's must take the initiative to ascertain if thinking skills should be accepted as fundamental facts for education or dismissed as futile faddism. As a profession EP's can drive the future in this sphere with research as the vehicle. Then and only then will the value of thinking skills instruction be established.

References

Adey, P. (2002). *Teacher Empowerment.* Paper presented at the 10th International Conference for Thinking. Harrogate, June 14th –16th, 2002.

Adey, P. (1992). The CASE results: implications for science teaching. *International Journal for Science Education,* 14, 137-146.

Adey, P. and Shayer, M. (1993). An exploration of long term far-transfer effects following an extended intervention program in the high school curriculum. *Cognition and Instruction,* 11 (1), 1-29.

Adey, P. and Shayer, M. (1994). *Really Raising Standards: Cognitive Intervention and academic achievement.* London: Routledge.

Alvino, J. (1990). A Glossary of Thinking-Skills Terms. *Learning,* 18 (6), 50.

Baird, J. R.(1986). Improving learning through enhanced metacognition; a classroom study. *European Journal of Science Education*, 8, 263-282.

Blagg, N. (1991). *Can we teach intelligence?* Hillsdale, NJ: Lawrence Erbaum Associates.

Bond, T. (2002). *Changing Minds and Empirical Evidence: Can we measure change?* Paper presented at the 10th International Conference of Thinking, Harrogate, June 14-16, 2002.

Chi, M,T,H. (1996). Constructing self explanations and scaffolded explanations in tutoring. *Applied Cognitive Psychology*, 10, 33-49.

Coles, M.J. (1993). Teaching Thinking: principles, problems and programmes. *Educational Psychology*, 13 (34), 333 – 344.

Connor, M J. (2002). *The Development of Thinking Skills.* Surrey Educational Psychology Service, April 2002.

Cotton, K. (1991). *Teaching Thinking Skills.* School Improvement Research Series(SIRS),North West Regional Educational Laboratory. hhtp://www.nwrel. org/scpd/sirs/6/cu11.html.

DeCorte, E. (1990). Towards powerful learning environments for the acquisition of problem solving skills. *European Journal of Psychology of Education*, 5, 5-19.

DeCorte, E. (2002). *Creating high-powered learning communities for thinking and problem solving.* Paper presented at the International Conference for Thinking, Harrogate, June 14 -16th, 2002.

Department for Education and Employment (1999). *From Thinking Skills to Thinking Classrooms; A review and evaluation of approaches for developing pupils thinking.* Research Report RR115. London, DfEE.

Department for Education and Employment (2000). *Educational Psychology Services (England): Current Role, Good Practice and Future Directions*. The Research Report. London, DfEE.

Feuerstein, R. and Y Hoffman, M.B. and Miller, R. (1980). *Instrumental Enrichment: An intervention for cognitive modifiability.* Baltimore, MD: University Park Press.

Fisher, A. (2002). *Assessing Thinking*. Paper presented at the International Conference of Thinking, Harrogate, June 14-16th, 2002.

Fitz-Gibbon, C.T. (2002). *What do schools need? How can Educational Psychologists contribute?* Keynote address at the Association of Educational Psychologists Annual Conference, Hinckley, October 24-26th, 2002.

Gardner, H. (1993). *Multiple Intelligencies: The theory in practice*. New York: Basic Books.

Gorodetsky, M., Barak J., Kaur, R,., Xamir J and Tirosh, R. (2002). *A Knowledge Constructing School.* Paper presented at the International Conference of Thinking, Harrogate, June 14-16th, 2002.

Hatcher, R. (2002). *The Business of Education*. Keynote address at the Association of Educational Psychologists Annual Conference, Hinckley, October 24-26th, 2002.

Higgins, S. (2002). *Teaching Thinking through pedagogical strategies in primary schools.* Paper presented at the International Conference of Thinking, Harrogate, June 14-16th, 2002.

MacKay, T. (2002). Discussion Paper- The future of Educational Psychology. *Educational Psychology in Practice*, 18 (3), 245-253.

McGuinness, C. (1993) Teaching Thinking: New Signs for Theories of cognition. *Educational Psychology*, 13 (3-4), 305-316.

McGuinness, C. (1999). *From Thinking Skills to Thinking Classroom: a review and evaluation of approaches for developing pupil's thinking.* DfEE Research Report, RR115.

McGuinness,C. (2000). ACTS A methodology for enhancing thinking skills across the curriculum. C.mcguinness@qub.ac.uk.

McGuinness, C. (2002). *Metacognition in Classrooms.* Paper presented at the 10th International Conference of Thinking, Harrogate, June 14 -16th, 2002.

McGuinness, C., Curry, C., Greer, B., Daly, P., and Salters, M. (1997). *Final Report on the ACTS project. Phase 2.* Belfast: Northern Ireland Council for Curriculum Examinations and Assessments.

McGuinness, C. and Nisbet, J. (1991). Teaching Thinking in Europe. *British Journal of Educational Psychology*, 61, 174-186.

Munro, J. (1999). Learning more about Learning Improves Teacher Effectiveness. *School Effectiveness and School Improvement*, 10 (2), 151-171.

Nisbet, J. (1993). The Thinking Curriculum. *Educational Psychology*, 13 (3-4), 281-290.

Pithers, R. T. and Soden, R. (2000). Critical Thinking in Education: A Review. *Educational Research* , 42 (3), 237-249

Portwood, M. (2002). *A metabolic basis for dyslexia, dyspraxia and attention deficit hyperactivity disorder.* Keynote Address at the Association of Educational Psychologists Annual Conference, Hinckley, October 24-26th, 2002.

Pramling, I. (1988). Developing children's thinking about their own learning. *British Journal of Educational Psychology*, 58, 266-278.

Rudestam, K.E. and Newton, R.R. (1992). *Surviving your Dissertation.* Newbury Park, California:Sage.

Salomon, G. and Perkins, D.N. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Journal of Educational Psychology*, 24, 113-142.

Savell, J.M., Twohig, P.T. and Rachford, D.L. (1986). Empirical Status of Feuersteins 'Instrumental Enrichment' (FIE) technique as a method of teaching thinking skills. *Review of Educational Research*, 56, 381-409.

Shayer, M. and Adey, P. (1993). Accelerating the development of Formal Thinking in Middle and High School Students IV: Three years after a two-year intervention. *Journal of Research in Science Teaching*, 30(4), 351-366.

Stewart, L. and Smardon, D. (2002). *Teachers and Learners participating in Critical Thinking for effective change*. Paper presented at the 10th International Conference of thinking, Harrogate, June 14th –16th. 2002.

Stoiber, K.C. and Kratochwill, T.R. (2000). Empirically supported interventions and School Psychology: conceptual and practical issues. *School Psychology Quarterly*, 15, 75-105.

Swartz, R. and Parks, S. (1994). *Infusing the teaching of critical and creative thinking into content instruction: A lesson design handbook for the elementary grades*. California: Critical Thinking Press and Software.

TES (2002). Thinking: the next big idea....*Times Educational Supplement.* June 14th 2002.

Topping, K. (2002). Paired Thinking. Developing Thinking Skills through Structured Interaction with Peers, parents and volunteers. In Reid, G. and Wearmouth, J . (ed) *Dyslexia and Literacy: Theory and Practice.* Chapter 18. John Wiley & Sons Ltd. 2002.

Watkins, C., Carnell, E., Lodge, C., Wagner, P., and Whalley C. (2001). Learning about Learning enhances performance. *The National School Improvement Networks Bulletin,* Spring 2001.

Wegerif, R., Higgins, S., Mercer, N., Dawes, L. and Sams, C.(2002).*Thinking Together: Research into Practice.* Paper presented at the 10th International Conference of Thinking. June 14th-16th,2002.

Wilson, V. (2000). *Can Thinking Skills be taught. A paper for discussion.* Scottish Council for Research in Education.

Appendix 1: An Example of the Questionnaire used for the Survey

Memo to: All Surrey EPs

SURREY
COUNTY COUNCIL

From: Jess Hudson

Tel:
Date: 25 June 2002

RE: THINKING SKILLS QUESTIONNAIRE

I am writing to request your assistance in completing a questionnaire regarding Thinking Skills. As you are aware, I am currently studying for a Doctorate at UCL and one of my professional assignments involves an evaluation of current practice in a specific area in an LEA context. I plan to focus on Thinking Skills and link it in with the current project Janet Wallace and I are working on. I would be most grateful if you could complete the attached questionnaire concerning current knowledge and practice in the domain of Thinking Skills. The information provided will give me an idea of the current situation in Surrey from which I can generate a contextual understanding.

Thank you for your assistance in this matter. If you have any questions please do not hesitate to contact me.

Please complete the questionnaire and return it to me in the SE Office by 31st July 2002.

Best wishes,

Jess

Att.

Questionnaire for Surrey Educational Psychologists regarding Thinking Skills

1. How much do you know about thinking skills ?

I----------I----------I----------I
1 (Nothing) 2 3 4 (A Great Deal)

2 How much do you know about the current research base in thinking skills ?

I----------I----------I----------I
1 (Nothing) 2 3 4 (A Great Deal)

3 What current techniques do you use to assess thinking skills ?

4 Which thinking skills do you target most in your work ?
Please tick the appropriate box.

Critical Thinking	☐
Decision Making	☐
Creative Thinking	☐
Problem Solving	☐
Metacognition	☐
Enquiry and Evaluation	☐
Other (please state)	☐

5 What strategies and interventions do you use to support or extend thinking skills ?

6 Which intervention packages of thinking skills are you aware of ?
Please tick the appropriate box.

Instrumental Enrichment (Feuerstein)	☐
Philosophy for Children (Lipman/Fisher)	☐
Subject Specific Thinking (Adey & Shayer)	☐
Infusion Methodology (McGuinness)	☐
Cognitive Approach (De Bono)	☐
Other (please state)	☐

7 Which intervention packages of thinking skills have you used ?
Please tick the appropriate box.

Instrumental Enrichment (Feuerstein)	☐
Philosophy for Children (Lipman/Fisher)	☐
Subject Specific Thinking (Adey & Shayer)	☐
Infusion Methodology (McGuinness)	☐
Cognitive Approach (De Bono)	☐
Other (please state)	☐

8 Which intervention packages of thinking skills have you seen implemented in schools in Surrey ? Please tick the appropriate box.

Instrumental Enrichment (Feuerstein)	☐
Philosophy for Children (Lipman/Fisher)	☐
Subject Specific Thinking (Adey & Shayer)	☐
Infusion Methodology (McGuinness)	☐
Cognitive Approach (De Bono)	☐
Other (please state)	☐

9 Would you like to be able to implement thinking skills packages within Surrey schools ? Please tick the appropriate box.

Yes	☐
No	☐

10 Any other comments:

Thank you for your help with this.

Please return this form to Jess Hudson, EPS at the SE Office.

Appendix 2: Limitations and Improvement Suggestions for Survey

The survey had restricted usefulness due to a variety of limitations. Firstly a limitation of the survey were the variables unaccounted for in the original analysis and design of the questionnaire. More information regarding demographic characteristics concerning EP's individual factors could have been useful. For example age, gender and years of practice may have had an important impact in shaping opinions and views on thinking skills. Likewise specialisms or senior roles may also have had an effect on the current knowledge and research base according to particular areas of responsibility or interest. Such information may have made analysis more detailed but also enlightening in its findings.

In terms of questionnaire design no consultation occurred with university advisors or EPS members in relation to the construction. Additionally certain items of the questionnaire were misleading. For example the question 'Would you like to implement thinking skills?' could have generated a misleading figure as not all EPs may be confident in providing INSET so would answer 'no' to this question despite being interested in thinking skills packages. Likewise the question which asked 'Do you know of any school using these packages?' created meaningless data as it would be influenced by firstly what EPs actually see, and secondly which schools that EP actually goes into. These items gave misleading responses, which may not have been representative. Moreover no measures of validity or reliability occurred. Similarly no pilots took place to ensure the smoothness of the survey process.

Other problems concerned the procedures utilised in data collection. Firstly the questionnaire was distributed at the end of the summer term. This may have meant EPs were too busy to complete it so only the most dedicated or interested EPs will have found time. Similarly EPs may have been taking leave at this point and the sample is therefore, limited and not generalisable to Surrey EPS as a whole. Findings cannot, therefore, be used as an accurate representation of Surrey EPs at present. The survey may also have been confounded due to the action of self-selection procedures. For example, perhaps it was only EPs who felt confident in their abilities or comfortable with self-appraisal that actually completed the questionnaire. The problem with such self-selection is that it is not known along which psychological dimensions it operates. Low response rates could be a result of the administration timing or that respondents felt responses could be used against them by employers. The reliance on self-report data is also an issue and it would have been appropriate to use a more objective external measure to account for the subjective information collected. For example schools rating the use of thinking skills techniques by EPs in context.

The survey could be improved upon in several ways to extend both its reliability and validity. A more balanced and representative sample could be selected to cover EPs thoroughly in a survey across services. Demographic characteristics could be ascertained and potential impact gauged. EPs could be interviewed and information of a more qualitative nature collected in area team meetings via focus groups. Likewise external measures from schools could provide a fuller picture of the actual psychology used relating to

thinking skills in schools. Finally liaison with Adrian Furnham, Professor of Psychology at UCL, regarding questionnaire design and analysis would have informed my skills and practice in this area.

Appendix 3: Table 1 ; An Overview of Thinking Skills Approaches

Generic Thinking Skills taught in context free situation. **Bolt- on Approach**	IE (Instrumental Enrichment) (Feuerstein, 1980; Blagg, 1991; Savell et al., 1986).	Thinking as Information Processing	Two by two pre-post design over 2 years. Issues of inappropriate measures and variables with no reasonable control.	Basic analysis and conclusions indicated no improvement in FIE trained groups in UK.	Clear conceptualisations and experimental design. Although some limitations of measures and variable categories.
Subject specific intervention to target thinking in certain curriculum areas. **Embedded Approach**	CASE (Cognitive Acceleration Through Science Education) (Adey & Shayer, 1993, 1994; Adey, 1997; Shayer, 1996).	Thinking as Sense Making	Two by two pre-post design over 2 years, extended later to 5 years. Concerns over lack of control group and extraneous variables unaccounted for.	Mixed results with effects unstable over time and gender. Conclusions not necessarily consistent with data.	Strong theoretical underpinnings with large scale and scope. Issues with research design.
Infusion methodology teaching thinking across the curriculum. **Infusion Approach**	ACTS (Activating Children's Thinking Skills) (McGuinness et al., 1997, 1999; Swartz & Parks, 1994).	Thinking Across Contexts	Exploratory study to investigate student and staff reactions to ACTS. Posttest of teachers only. No clear research design with adequate controls.	Perceived positive impact, however, measures lacked reliability and validity. Limitations were identified. Improvements highlighted.	Useful basis for follow-up research. Key difficulties relate to research design and confounds.

Professional Practice Assignment 2: Nurture Groups: Evidence of Effectiveness

Abstract

Nurture groups are an intervention, which have been used during the last forty years and have been the focus of renewed interest in the current educational climate (Connor, 2001; O'Connor & Colwell, 2002). This paper details the theoretical origins of the approach and critiques current evidence of effectiveness. Such evaluation reveals that, although benefits exist, considerable scientific study is warranted to establish the link between nurture groups and long term improvements for children. The current importance of resolving earlier research limitations and examining practice issues are highlighted. Future developments for Educational Psychology are illustrated through research into attachment theory and alternative interventions.

Section 1: Aims and Scope of Assignment

As Cooper and colleagues (2001) point *out "interest in Nurture Groups continues to grow"* (p. 160). Indeed *"researchers, policy makers and practitioners continue to be interested in the impact of Nurture Groups"* (O'Connor & Colwell, 2002) (p. 96). A resurgence of attention on nurture groups has been stimulated by the emphasis on promoting inclusion in schools whilst reducing exclusions (Bennathan, 1997; Connor, 2001; Cooper, 2003). Shearman (2003) comments: *"current educational policy is oriented towards including children with special educational needs in mainstream*

classrooms" (p. 53). Nurture Groups offer a *"nurture-based practice as one approach to the inclusion of pupils with social, emotional and behavioural difficulties"* (Cooper et al., 2001) (p. 160).

At the other end of the inclusion spectrum, there has been an unprecedented increase in UK school exclusions in the 1990s which, led to a variety of interventions for emotional and behavioural difficulties being investigated (DfEE, 1999a; 1999b). In particular *"the rise in numbers of pupils excluded from school has provoked considerable concern amongst policymakers and professionals in education and beyond"* (Head et al., 2003) (p. 33). Renewed interest in Nurture groups has, therefore, been stimulated, in the hope that they could provide an innovative approach to the inclusion of children with emotional and behavioural difficulties (Cooper & Lovey, 1999; Iszatt & Wasikewska, 1997). Furthermore there was endorsement of the model in the DfEE (1997) Green Paper.

In this professional practice assignment, the aim is to provide a critical overview of the evidence from Nurture group practice to ascertain effectiveness. Such a critique will then form the basis from which implications for Educational Psychology practice can be explored. Early intervention is particularly important for practising EPs as it is listed as one of their core functions (DfEE, July 2000).

Definitions

Nurture Groups were first operationalised by Marjorie Boxall, an Educational Psychologist, in the 1970's. The groups *"take the form of discrete classes, located in mainstream primary school, for approximately 12 students, staffed by a teacher and specially trained Teaching Assistant"* (Cooper et al., 2001) (p.160). The core concept underlying Nurture Groups is 'attachment theory', which focuses on the formation of an ongoing relationship between child and caregiver during the early years (Bennathan, 1997; Connor, 2001). The nature and quality of such relationships were predicted to affect both the short and long term psychological development of children (Bowlby, 1951). The nurture group intervention aims to compensate for missed critical learning opportunities and nurturing in children's early development which, are likely to generate emotional and behavioural difficulties indicative of earlier egocentric behaviour (Cooper & Lovey, 1999). The Nurture Group experience aims to help children *"progress from this state to the level of social competence that is required in the standard infants' school classroom"* (Cooper et al., 2001) (p. 160). Nurture groups provide an environment *"where young children can re-experience early nurturing"* (Iszatt & Wasilewska, 1997) (p. 64). Students therefore receive input *"both in relation to the formal curriculum and in terms of their social and emotional functioning"*(Cooper & Lovey, 1999) (p. 124). Gradual reintegration and full time participation into the attached mainstream class is built into the model (Cooper et al., 2001).

Overview

The purpose of this assignment is to examine the evidence for Nurture Group effectiveness and to outline future directions for EP practice. Initially the paper will detail the relevance of Nurture Groups at a local and national level. A critique of the existing research will follow and this will be used to inform the discussion into future developments in Educational Psychology.

Section 2 : Practice and Context

The impetus for investigating Nurture Group effectiveness exists at a number of different levels. Practice at a local level will be illustrated followed by a national focus. In both of these, summaries of current practice will be provided and dimensions for further exploration identified.

Local Context

In the last year, Education Children's Services in a large shire county authority have invested considerable resources in investigating and implementing Early Years interventions. This venture is targeted in both the Education and Behaviour Development plans as warranting exploration in the context of increasing levels of emotional and behaviour difficulties in school (Cooper, 2003; Shearman, 2003). As part of this countywide initiative, a pilot project was established to examine the effectiveness of a Nurture Group and to inform discussions about using the approach on a countywide level.

The Nurture Group pilot was established in a school with a significant population of students with emotional and behavioural difficulties. In Spring 2002 a steering committee was set up to plan, supervise, manage and monitor the new provision. Initial staff training occurred in Summer 2002 and baseline (pre) assessment data from children, parents and staff were collected. In September 2002 the Nurture Group commenced with regular follow-up and monitoring from the Steering Committee. Post-baseline assessment is due in the summer term 2003.

At a local level, this pilot represents a laudable and constructive LEA evaluation of the efficacy of Nurture Groups, but viewed from a psychological perspective there are obvious drawbacks, which are likely to limit the value of the outcomes. Unfortunately the lack of any control group in the evaluation design means a range of extraneous variables are likely to affect the outcomes. Likewise, systematic bias in the sample limits conclusions that can be drawn. Finally, the lack of standardised assessment measures will affect the study's credibility. In short, at a local level there is a mismatch between an espoused rigorous study and the implemented evaluation where methodological limitations exist. This is cause for concern given the potential impact of the evaluation on LEA decisions on future and extended Nurture Group implementation.

National Context

On a national level the rationale for this assignment originates from three levels: from society itself, schools as organisations and psychology as a discipline. Each of these areas will now be examined.

Society as a Whole

The relevance of Nurture Group intervention at a societal level has arisen as a result of current issues in education. As Cooper (2003) details *"the topic of including students with Special Educational Needs in mainstream schools is itself of current interest to educators throughout the world"* (p.5). Indeed *"In Britain at the moment there is a strong national and local policy commitment to the notion of inclusion"* (Lloyd et al., 2003) (p.87). This is evident in government legislation and education protocols alike (CSIE, 2000; DfEE, 1997; DfEE, 1998). In addition to this there are the increasing incidence rates of emotional, behavioural difficulties and related school exclusion (DfEE, 1999a, 1999b; Hayden, 1997). As Bishop and Swain (2000a) highlight *"it has become evident that, over the last few years, exclusions from schools have increasingly included younger children"* (p.18). Indeed it is acknowledged universally that exclusion is a critical issue (Cooper, 2003; Geddes, 2003; Head et al., 2003; Shearman, 2003).

In relation to this Nurture Groups offer an attractive approach which addresses both inclusion and exclusion. As a result *"recently, the UK government has identified NGs as a promising intervention for children with EBD"* (Cooper & Lovey, 1999) (p. 123). Nurture Groups have been regarded

"as an effective early method of identification and intervention for some pupils with emotional and behavioural difficulties" (Doyle, 2001) (p. 126) and their popularity is evident in government papers (DfEE, 1997). Nurture Groups not only tackle a range of issues at an early stage, but the approach encompasses the crucial role parents have on children's development (Bishop & Swain, 2000a, 2000b; Deater-Deckard, 2000; DfEE, 1997, 1998; Marvin et al., 2002). The extent to which Nurture Groups resolves these educational issues is obviously a big question.

Schools as Organisations

At a school level there is increasing evidence as to the positive impact of Nurture Group principles and practice to the whole organisation (Bennathan & Boxall, 2001; Connor, 2001; Doyle, 2001; Lucas, 1999). It is reported that *"teaching and learning becomes more effective for all children"* (Lucas, 1999) (p. 14). Moreover, it has been claimed that the approach acts as *"a positive alternative experience for pupils at risk of exclusion"* (Doyle, 2001) (p.131). Indeed adopting the Nurture Group philosophy is said to provide *"an exceptional opportunity to evolve as a school"* (Doyle, 2001) (p. 131). As O'Connor and Cowell (2002) argue, work *"begins to suggest a need for a whole school nurturing approach"* (p. 99). The key questions which emerge, would seem to be whether such promising effects exist and if so, what are the critical factors involved?

Psychology as a Profession

The theoretical underpinnings of Nurture Groups are of considerable professional interest to Educational Psychologists given their origins in

Attachment Theory. This can inform both assessment and intervention in the early years, a core function for EP's (DfEE, July 2000). As Steele (2002) suggests attachment theory *is "the most powerful contemporary account of social and emotional development available to science"* (p. 518). Indeed Turner (2003) describes how attachment theory is a theory of child development and has implications for the teaching and learning of children with emotional and behavioural difficulties. Attachment Theory (Bowlby, 1953; 1969) stated that infants engage in proximity-seeking behaviour when sensing discomfort and attempt to get closer to the attachment figure (primary carer). On the basis of the caregiver reactions the child constructs internal working models. As Glaser (2000) highlights *"these models are beliefs by the child about herself or himself and predictions about how he or she will be treated by others"* (p. 102). The ongoing implications of this are delineated by Rutter (1995) who details how *"children's experiences of interpersonal relationships are crucial to their psychological development"* (p. 550).

Attachment typologies have been identified (Ainsworth, 1967; Ainsworth et al., 1978) and attachment disorders of childhood have been included in classification systems of both DSM-IV and ICD-10 (Rutter, 1995). Nurture Groups also link into *"current theoretical understandings, which indicate close associations with social competence... and social and cognitive models of academic self regulation"* (Cooper & Lovey, 1999) (p. 123). All these theoretical foundations make Nurture Groups an important development for EPs to be involved in, to extend existing conceptualisations.

The impact of attachment theory on later development has been highlighted in that *"early attachments might play a role in the genesis of later psychopathology"* (Rutter, 1995) (p.531). Indeed work combining developmental psychology and neuropsychology has demonstrated the impact of attachment on brain growth and functioning (Balbernie, 1999; Glaser, 2000; Greenfield, 1997; Robertson, 1999; Svanberg, 1998). Svanberg (1998) argues *"bringing up children who are not securely attached is like burdening them with a traumatic psychological legacy"* (p. 128) (see Maher, 2002 for a discussion of this point). It appears that attachment is a life long concept which can permeate one's existence in a variety of ways (Turner, 2003).

Recent research does support the claim that prolonged disruption to attachment relationships are closely linked to serious emotional disturbance in children (Kobak et al., 2001). Rutter (1995) also applies attachment concepts to conduct problems and patterns of social withdrawal and concludes, *"further research will be needed to show how far attachment concepts are useful in gaining a better understanding of the mechanisms involved in the genesis of these disorders"* (p.560).

The concepts from attachment theory have also influenced intervention models as Marvin et al (2002) note "*over the past few years the number of systematic intervention programmes that are at least partially driven by attachment theory and research have been increasing*" (p. 108). This is an area ripe for development and this assignment intends to illustrate the crucial

role EPs could play in linking attachment theory to interventions (c.f. Mackay, 2002).

Summary

Considerable agreement exists at both a local and national level as to the current educational issue of inclusion of children with emotional and behavioural difficulties. Likewise, both fields of practice regard Nurture Groups as a potential intervention to resolve these concerns. As a result of this increasing interest there is a demand to examine the impact of Nurture Groups. Evidence of successful practice is in place as research projects in the local and national context are apparent. The constraint of how rigorously these claims have been established, however, remains.

The key controversy in both areas of practice is the issue of how scientifically Nurture Groups are being evaluated. At present there is a mismatch between espoused replicable study of Nurture Groups and the actuality of real life research. In both domains a need exists to determine the true impact of Nurture Groups. An aim of this assignment is to ascertain the evidence of effectiveness to date. Additionally a range of other dimensions need exploration, this paper aims to delineate these important variables, which have hereto been ignored in practice.

Section 3 : Psychological Theory and Research

The literature search used bibliographic databases of PsychInfo and ERIC as recommended by Ramchandani et al (2001). The key search terms used were "Nurture", "Nurture Groups" and "Attachment". Manual searches from recent journals occurred in conjunction with ancestral searches. In addition websites were used including:

- the Nurture Group Network (www.nurturegroups.org),
- the International Attachment Network (www.attachmentnetwork.org)
- the Attachment and Human Development Journal (www.tandf.co.uk/journal/titles/14616734.hton).

The review aims to critique the literature surrounding the evidence of effectiveness of Nurture Groups in the United Kingdom. The evidence to be examined occurred with UK samples and was selected on this basis. Likewise the research was reported as the existing empirical foundations (Connor, 2001; Cooper et al., 2001; Nurture Group Network). The review is divided into five sections using the critical evaluation checklist from Rudestam and Newton (1992). Initially the conceptualisations and theoretical frameworks of the literature are explored. Following on from this the research design is critiqued alongside the results and discussion. A final section details the major themes and controversies in a summary.

1. *Conceptualisation*

Studies in the Nurture Group literature identified the major issue of investigation as the need to ascertain the impact of the intervention. Research from Bennathan (1997); Cooper et al (2001) and Iszatt and Wasilewska (1997) attempted to study the effectiveness of Nurture groups in addressing children's emotional and educational needs. Whilst work from Cooper and Lovey (1999) and Bishop and Swain (2000a, 2000b) explored the *perceptions* of outcomes associated with Nurture Groups. All studies identified an issue for investigation accurately and delineated it accordingly. Likewise all studies build on from one another and are original in focussing on different aspects of evaluation.

The majority of the research contextualised the intervention by placing it as a response to the key issues in society and the prevailing educational climate. In particular the increasing concern over children with emotional and behavioural difficulties were highlighted alongside the associated rates of exclusion in Bennathan (1997); Bishop and Swain (2000a, 2000b) and Cooper et al (2001). Additionally the idea of inclusive education and the need to meet all children's needs in mainstream settings was cited as another justification in Bennathan (1997); Bishop and Swain (2000a, 2000b) and Cooper et al (2001). In contrast to this, research by Cooper and Lovey (1999) did not provide any rationale as to why perceptions were being explored. Similarly the paper by Iszatt and Wasilewska (1997) failed to provide justification for this type of research.

Overall the major concepts were clearly defined in the literature with the origins of Nurture Groups detailed and practical implications described. In terms of theoretical underpinnings, however, research evidence varied. Some studies clearly delineated the theory base such as Cooper and Lovey (1999) and Cooper et al (2001); whilst others only mentioned them briefly in Bennathan (1997); Bishop and Swain (2000a, 2000b) and Iszatt and Wasilewska (1997). This highlights a restricted theoretical explanation from the latter studies impacting on credibility.

2. Theoretical Framework and Hypotheses

A clear theoretical framework in relation to attachment theory (Bowlby, 1969) and Nurture Group philosophy was delineated in recent nurture group research from Cooper and Lovey (1999); Cooper et al (2001) and O'Connor and Colwell (2002). In these studies the manner in which nurture groups are designed to meet the needs of children in relation to attachment are described. Earlier pieces of work, however, made little reference to the underlying theory base. Interestingly, no research papers make reference to attachment typologies (Ainsworth et al., 1978) or general developmental theories. This seems to limit thinking in research to one theory base, that of Bowlby (1969), which may be restrictive in considering all the potential factors inherent in early relationship formation.

Broad aims and objectives were present in all the literature reviewed and relate principally to ascertaining the effectiveness of Nurture Groups.

Despite this, however, a number of studies had no clearly stated research questions with a rationale namely Bennathan (1997); Cooper and Lovey (1999) and Iszatt and Wasilewska (1997). In these articles the absence of any formulated research questions reduces credibility and usefulness. Fortunately recent work used focused research questions, which were operationalised into key areas of investigation as in Cooper et al (2001) and in Bishop and Swain (2000a, 2000b). Although, even in these papers the selection of research questions is not justified. Hypotheses were absent from all of the above research projects and as a result, the relationships amongst variables are unclear with key variables not explicit. Replication is, therefore, impossible and the integrity of studies undermined.

3. Research Design

Unclear or unspecified research design is inherent in the main studies of Nurture Group research. Indeed, Bennathan's (1997) work stated no design type preferring instead a casework style analysis. Likewise, work from Bishop and Swain (2000a, 2000b) used qualitative evaluation, but with no detail of procedures utilized, while Cooper and Lovey (1999) aimed to sample views of practitioners using a questionnaire, again with insufficient evidence of design. Finally, Iszatt and Wasilewska (1997) used a variety of evidence with a confusion of experimental types. The only clear designs are present in the longitudinal follow-up of O'Connor and Colwell (2002) and the intervention evaluation study of Cooper et al (2001).

Such a general absence of research design is inappropriate to this field of study, given that all these papers were attempting to gauge the impact of Nurture Groups. It, therefore, seems appropriate that some form of comparison with a control is necessary. Additionally, all studies have problematic descriptions of the Nurture Group intervention with no detail on time or duration. This failure in providing sufficient information on the intervention makes replication difficult. Similarly it will be unclear which variables are making an impact. In future, a clear procedure and design would ameliorate these issues as in recent work on attachment (Deater-Deckard, 2000; Kobak et al., 2001).

In terms of extraneous variables no study adequately controlled for the range of confounding factors. In the 1997 papers by Bennathan and Iszatt and Wasilewska, variables were not identified or controlled for. Later pieces evidenced high levels of systematic bias with work from Cooper and Lovey (1999) using only practitioners committed to Nurture Groups or Bishop and Swain (2000a) which used a self-selecting school dedicated to the approach. These may have led to skewed samples. In the one study, which used a control group (Cooper et al., 2001) limitations were still apparent with selection bias and systematic differences between comparison groups. To enhance scientific rigour, future work would need to use randomised control groups, with researchers blind as to group allocation. Moreover matched samples to eliminate systematic differences could be helpful. Finally, detailing demographic characteristics in relation to the sample would be critical in minimising confounds. As Bretherton (2000) points out mediating

variables such as emotional availability or maternal sensitivity are ignored. Likewise the extent to which Nurture Groups are a mediating mechanism are not examined (O'Connor et al., 1998).

In relation to the samples used in the research there is great variation. Initial work by Bennathan (1997) and Iszatt and Wasilswska (1997) provided no detail on samples in terms of demographic characteristics, recruitment, and inclusion statistics. This has implications for the applicability of findings to the national population. Later studies demonstrated how participants were selected but, with substantial recruitment bias evident. Indeed in Cooper and Lovey (1999) and Bishop and Swain (2000a) the views of professionals committed to Nurture Group interventions were used. Likewise O'Connor and Colwell's (2002) sample selected those who had sustained experience of Nurture Groups. All indicate skewed samples in favour of the intervention.

In the one study, which does employ a control group (Cooper et al., 2001) the sample is not analysed to gauge baseline differences between the groups, which may invalidate differences ascribed to the intervention. An obvious improvement would be to list the important features of groups so a comparison can be made. Finally in all of the research sample size is not calculated. In future, work will need to determine the power of the sample to detect a statistically significant effect.

Details of measurement devices used in the studies discussed are limited. In quantitative studies no indication of validation, reliability or validity statistics

are evident including Bennathan (1997); Cooper et al (2001); Iszatt and Wasilewska (1997) and O'Connor and Colwell (2002). In relation to the qualitative studies, although questionnaires and/or semi-structured interviews were identified, no detail is provided on construction or external validation. This is in direct contrast to standardised evidence based assessment techniques used in attachment research (Marvin et al., 2002). Moreover all the Nurture Group studies use a limited number of measurement devices with no triangulation of evidence. This again is in juxtaposition to attachment research where multiple informants and assessments are utilised (Deater-Deckard, 2000; Kobak et al., 2001; Marvin et al., 2002). At present Nurture Group study is limited by the quality, number and range of measurement devices used. For example, all Nurture Group research relied on the Boxall Profile (Bennathan & Boxall, 1998) to gauge children's outcomes. A range of criticisms have been levelled at this device (Connor, 2001; O'Connor et al., 1998). In future individualised assessment of the attachment type of the child could be explored (Marvin et al., 2002). Likewise measures to ascertain a child's perception of attachment could be used (Green et al., 2000). Adult measurement devices are similarly limited, but could be extended to combine self-report and interview measures (Bartholomew & Moretti, 2002). The over- reliance on parental self-report would also need addressing as detailed by Belsky and Fearon (2002) with alternatives of Emotional Availability Scales examined (Bretherton, 2000).

The study size in the majority of research was not large enough to have a high chance of detecting a significant impact in Bennathan (1997); Bishop

and Swain (2000a); Cooper and Lovey (1999) and Iszatt and Wasilewska (1997). In addition, all of these papers had insufficient follow up and limited duration, which restricts credibility. The two longitudinal pieces by Cooper et al (2001) and O'Connor and Colwell (2002) have a sustained time period and follow-up. Credibility of both, however, needs to be strengthened by the use of adequately matched control groups.

4. Results and Discussion

Studies investigating Nurture Group impact tended not to set the scene accurately in relation to results. Indeed only the papers from Iszatt and Wasilewska (1997) and Cooper et al (2001) detailed the time and point at which assessments were made. Furthermore the only study to report on the detail of the sample in terms of variables was O'Connor and Colwell's (2002) piece. No study in the research literature used any statistical adjustment for baseline differences in key variables.

Such a factor may mean confidence and probability values will be misleading (Greenhalgh, 2001). Even in the only piece to use control groups (Cooper et al., 2001) no statistical analysis was made to determine the extent to which groups were comparable. In terms of data types a range was evident yet no paper provided a rationale for selection. In fact some studies used a number of data types including Bishop and Swain (2000a) and Iszatt and Wasilewska (1997) with no explanation.

The majority of studies made no reference to statistical techniques or methods of analysis. In Bennathan's work (1997) and Iszatt and Wasilewska (1997) some data are presented in tables, but only uses percentages. Of the more qualitative data no detail is provided on the method of systematic analysis. Instead comments are reported on in groups or descriptors in both Bishop and Swain (2000a, 2000b) and Cooper and Lovey (1999). In this manner there is a high degree of systematic bias evident with causality assumed. No negative comments are evident and detail on responses omitted. The only two pieces of research attempting to use any form of statistical analysis are Cooper et al (2001) and O'Connor and Colwell (2002). Both pieces calculate mean differences in clear tabular form with probability values and confidence intervals. Despite this, however, these values are not interpreted nor reflected in the paper's text. Similarly outliers are not addressed and no data is presented as to the control group. As a result assumptions are evident in the nature and direction of causality.

All the studies in the research literature inadequately manage and account for control variables in the data analysis. As Belsky and Fearon (2002) indicate, measures of refusal rates are important yet these are not detailed in any paper. The extent to which Nurture Groups are a mediating variable is, likewise, unaddressed (O'Connor et al., 1998). Furthermore, no study demonstrates any consideration of the impact of other variables from the role of parenting (Deater-Deckard, 2000), to child characteristics (Belsky & Fearon, 2002) or the time frame involved (Deater-Deckard, 2000). These

potentially relevant factors are unaccounted for in the results, which represents a significant omission.

With regards to the conclusions drawn, a number of studies have discussions which, are consistent with the results including Bennathan (1997); Bishop and Swain (2000a, 2000b) and Cooper et al (2001). However, in the examples cited there is a tendency to over-generalise from the results, for example in detailing the positive impact of Nurture Groups on whole schools when the study focused on a small group. Moreover, there is wide extrapolation from results as to positive outcomes in all domains,that is not evidenced by statistical analysis, as in Bennathan (1997); Bishop and Swain, (2000a, 2000b) and Cooper et al (2001). The studies of Iszatt and Wasilewska (1997) and Cooper and Lovey (1999) draw conclusions, which are not explicit in the evidence. In all of the studies mentioned alternative conclusions consistent with the data are not discussed. As such the majority of studies assume a significant difference is due to Nurture Groups. Given the limited power of studies this may be a Type 1 error, when a significant difference is attributable to sampling error rather than a true difference (Greenhalgh, 2001). The only research, which delineates between long and short-term impact and alternative conclusions, is that of O'Connor and Colwell (2002).

In terms of detailing theoretical and practical implications, the studies on nurture groups are varied. In the work of Cooper and Lovey (1999) and Cooper et al (2001) no discussion on either of these issues is available. In

other papers, the practicalities of the work are detailed with Bennathan (1997) requesting more *"systematic evaluation"* and O'Connor and Colwell (2002) the professional support of teachers. The only study to consider, the more controversial issues related to Nurture Groups as a concept is Bishop and Swain's piece (2000a). In this paper the deficit view of the family in the "transplant" model of teacher-parent partnership is highlighted.

Limitations in the research are not identified in the work of Bennathan (1997); Bishop and Swain (2000a) and Cooper and Lovey (1999). This restricts the usefulness of the studies in providing a foundation for continued work in the field. The paper by Cooper et al (2001) identifies limitations appropriately and suggests improvements to resolve these. Likewise the O'Connor and Colwell (2002) piece highlights problems inherent in measurement devices used and the atypical sample. Iszatt and Wasilewska's (1997) research also refers to limitations such as the *"inconsistencies of applying Code of Practice criteria and the high mobility of the population"* (p.69). No studies, however, recognise the role of confounding variables.

5. Summary of Psychological Literature

Current psychological research suggests Nurture Groups have some positive impact on students and schools (Connor, 2001). There are, however, several issues with the adequacy of studies to date, which require improvement.

Overall, the research into Nurture Group effectiveness has built upon itself and improved. Initial work by Bennathan (1997) and Iszatt and Wasilewska (1997) offered a clear basis and rationale for the approach. As research studies, however, they have limited value with no experimental design making replication difficult and undermining conclusions drawn. Later research of a more qualitative nature from Bishop and Swain (2000a) and Cooper and Lovey (1999) provided an insight into the perceptions of the personnel involved in Nurture Groups which was original. Both papers, however, were riddled with methodological flaws with non-validated measurement devices and samples with systematic bias skewing results. Recent work has been of a higher adequacy in terms of experimental design and credibility. O'Connor and Colwell's study (2002) examined children in Nurture Groups over a two year period to ascertain the maintenance of change. In this study, however, no comparison group was utilized. The most credible study to date is Cooper et al (2001) which offers the first scientifically rigorous attempt to evaluate Nurture Groups in the field. It has an experimental design with matched controls occurring over a significant time period with substantial sample. Even in this piece, however, measurement issues will limit its worth in detecting difference attributable to Nurture Groups. Research has developed in this domain but, considerable issues still exist with regard to experimental design.

The main controversies that remain within the present literature relate to research limitations and unresolved issues. As previously stated, the key drawback of papers pertain to issues of research design. Firstly the majority

of research lacked clear comparison groups by which to evaluate Nurture Group effectiveness. Likewise the measurements used were not detailed in terms of validity, reliability or rationale for selection. The size and scope of the studies also limited credibility. Furthermore the practical implementation of the approach at an organisational level were not examined. A number of unresolved issues were similarly apparent linking to both practical and theoretical implications. The role of attachment was not thoroughly investigated within the Nurture Group context with the impact of infant's internal working models (Rutter, 1995) and time span ignored (Belsky & Fearon, 2001; Kobak et al., 2001). Indeed, a range of variables were unaccounted for from the child (Rutter, 1995) to the environment and genetics (Deater-Deckard, 2000; O'Connor et al., 1998). From a theoretical perspective the development of attachment theory in relation to Nurture Groups was unaddresssed. Moreover, the practicalities of interventions aimed to ameliorate attachment difficulties were not explored nor alternatives examined.

To conclude, the research, although indicative of a positive impact on students is limited by a range of methodological and measurement flaws. As O'Connor and Colwell (2002) argue there is a *"call for more in-depth, longitudinal research into Nurture Groups"* (p. 96) with studies of a more scientifically rigorous nature required. This will need to include the resolution of earlier limitations with clear research design and measurement selection. Exploration of unresolved issues and a discussion of both the theoretical and practical impact of the intervention will, likewise, be warranted.

Section 4 : Integration of Theory, Research and Practice

This section outlines the manner in which current research and practice both overlap and conflict. The shortfall between these areas will be explored and implications for action delineated. In particular the role Educational Psychology may play will be detailed in resolving earlier limitations, examining issues in practice and pursuing future research directions.

Research and Practice Issues

Some overlap exists between research and practice along key dimensions in terms of both the rationale and theoretical basis underlying Nurture Groups. Both practice and research, evidence the need to reduce exclusion rates nationally and locally (Bennathan, 1997; Bishop & Swain, 2000a; Cooper, 2003; Iszatt & Wasileskwa, 1997; Shearman, 2003). In particular *"the rising tide of social, emotional and behavioural difficulties (SEBD) among young children in the early years of schooling"* was identified (Cooper et al., 2001) (p.160). This issue was clear in both practice (Cooper, 2003; DfEE, 1999a, 1999b; Lloyd et al., 2003) and research (Bennathan, 1997; Bishop & Swain 2000a; Cooper et al., 2001; Iszatt & Wasilewska, 1997). Likewise the inclusion of these students into the mainstream setting was a matter of overlap with the need *"to develop more inclusive approaches to young people with social, emotional and behavioural difficulties"* (Head et al., 2003) (p. 33). Nurture Groups offer a potential resolution to both these problems (Bennathan, 1997; Bishop & Swain, 2000a; Cooper & Lovey, 1999; Doyle, 2001). As Cooper et al (2001) highlight *"practitioners are enthusiastically*

exploring nurture-based practice as one approach to the inclusion of pupils with social emotional and behavioural difficulties" (p. 160). This also links in to the growing need for early intervention based on resolving attachment difficulties (Deater-Deckard, 2000; Marvin et al., 2002). Finally, the theoretical basis of Nurture Groups in attachment theory is evident across research and practice (Connor, 2001).

Although there is agreement as to the potential of Nurture Groups as an intervention, the areas of practice and research conflict as to proven effectiveness. Whilst practice demands a scientifically rigorous analysis of Nurture Groups, the research fails to provide this. As Cooper et al (2001) point out *"evidence from systematic evaluation studies is rare"* (p. 161). Indeed, the review of the psychological literature indicates research manifest with methodological flaws alongside a range of unresolved issues. It appears a shortfall exists between what practice requires and what research has hereto provided. Such an imbalance will need re-addressing so research is able to answer the questions set in practice.

Local Level – Short Term Action to Resolve Limitations

A review of the literature and local level research highlighted a number of confounding variables implicit in research designs. In terms of measurements the majority of studies used only the Boxall profile to ascertain student outcome. This device as Doyle (2001) indicates is *"too broad"* and a need exists for a more *"specific, quantitative assessment tool"* (p. 127). Future work at a local level could develop and extend current

measurement devices to tap into attachment. As Bartholomew and Moretti (2002) argue *"more systematic research is required to investigate the predictive validity of different methods of assessing attachment"* (p. 164). In the local pilot study alternative measurements could be used to explore the child's level of attachment. Examples such as Green et al.,'s (2000) toy scenario could be used or the pre-school attachment classification scale (Marvin et al., 2002). At an adult level (an area unaddressed to date) techniques such as the Emotional Availability Scales (Bretherton, 2000) or Adult Attachment Interview (George et al., 1985) could be investigated. The use and inclusion of these instruments in practice would add to the assessment and understanding of difficulties. These areas of exploration would also inform the theory base with regard to attachment (Bartholomew & Moretti, 2002).

Another key limitation, which will need overcoming at a local level, is the research design and size of projects. Previous work used inadequate controls, which received no alternative intervention in comparison to the Nurture Group. In future practice, research designs could use a waiting list, attention-placebo or alternative intervention control group. Samples could be larger and students matched to represent the national population. Pre and post measurement over more substantial time periods could occur to determine the optimum duration for intervention.

Furthermore the scope and impact of Nurture Groups could be expanded to a whole school level. Research does indicate some positive effect to schools

as organisations (Bennathan, 1997; Bennathan & Boxall, 1998; Lucas, 1999), yet as Cooper et al (2001) point out *"this is still unresearched aspect of Nurture Groups"* (p. 161). At a practice level issues of implementation could be explored on a whole school basis to tackle EBD (Bennathan, 1997; Bennathan & Boxall, 1998; Bishop & Swain, 2000b; Lloyd et al., 2003; Lucas, 1997). Consultation with schools regarding attachment and nurturing principles, which are pertinent to the learning environment, could occur. As Bennathan and Boxall (2001) indicate there is a need *"to establish a nurturing ethos throughout the school as a solid basis for education"* (p.64). The extent to which Nurture Groups act as an inclusive procedure could be examined (Bishop & Swain, 2000a; Connor, 2001). Similarly issues with the "transplant model" of teacher expertise being transferred to parents in Nurture Groups could be addressed (Op.Cit). These issues could be tackled in local practice, with pilot projects to extend the size and scope of studies whilst also ascertaining practical wider impact. Educational Psychologists could construct such research at a small scale to ascertain effectiveness over time. In practice this would involve supporting Nurture Group schools in planning and executing evaluative research.

Educational Psychologists may need to look fundamentally at practice, to gauge if the education system needs to change its structures and relationships to accommodate for children with disordered attachments (Turner, 2003). In particular, examining whether developing attachments to significant adults in schools can assist children's development (Turner, 2003).

National Level – Medium Term Action to Examine Unresolved Issues

A number of key controversies were left unexamined in the literature, which could be readdressed by targeted research. Firstly the role and remit of attachment as a concept was left untouched. Research at a national level could explore the intricacies of such a phenomena (Steele, 2002; Turner, 2003). In particular infant's internal working models (Bowlby, 1980) in terms of attachment could be ascertained. As Belsky and Fearon (2002) argue *"the specific role of the IWM has been less subject to empirical scrutiny"* (p.362). These could be examined in relation to Nurture Group impact (Belsky & Fearon, 2002; Green et al., 2000; Rutter, 1995) using scales to measure internal working models of attachment (Main in Goldberg et al., 1995). Likewise, the impact of attachment disruption could be examined (Kobak et al., 2001). This could extend research to children across the age range for as Kobak et al (2001) highlight *"much less is known about how major disruptions in attachment relationships influence older children's adjustment"* (p. 244).

The impact of time on attachment disorders and mediating experiences could likewise be investigated (Belsky & Fearon, 2002; Kobak et al., 2001). As Belsky and Fearon (2002) argue *"debates about the time-span effects of early attachment insecurity on later psychological functioning and questions about modes of transmission have occupied a prominent role in theorizing about attachment"* (p. 379). These issues could be explored within Educational Psychology to ascertain attachment and how it could be moderated by later experiences to ameliorate potential negative effects.

Furthermore the role and impact of attachment at different stages of development could be gauged to discover issues arising across childhood (Greenberg, 1999). This would theoretically develop the concept of attachment alongside informing developmental and educational psychology in terms of assessment and intervention over time.

Another realm for further investigation are the variables unaccounted for in the research literature, these range from child factors to the impact of the environment. In terms of child characteristics there is the role of temperament and how this *"might influence the qualities of attachment"* (Rutter, 1995) (p. 553). Medium term research could begin to ascertain the interaction between children's characteristics and attachment formation, tracking potential differences over time. Connections to children's behaviour control systems could be explored, linking to research suggesting neurobiological responses to attachment types (Rutter, 1995). Investigation into this could link developmental and neuropsychology together in determining the origins and interventions for attachment disorders. This has implications for Nurture Group research, as these factors would need to be controlled for by using multiple control and experimental groups. At a practice level implications from neuropsychology (Balbernie, 1999; Robertson, 1999) could be disseminated by EP's to support the management of children in schools.

The impact of different relationships on attachment have been omitted in the research literature. As Rutter (1995) highlights *"there have been instances in*

which attachment concepts have been over-generalized or misinterpreted in a naïve and simplistic fashion" (p. 566). In Nurture Group research this view has tended to prevail and in future other relationships must be considered alongside attachment. Such work could explore attachment alongside maternal sensitivity, emotional availability, peer and sibling relationships. As Rutter (1995) states *"what is needed now is a bringing together of attachment concepts and other formulations of relationships so that each may profit from the contributions of the other"* (p. 566).

Other variables such as demographic characteristics ignored in the Nurture Group research will need exploration (Kobak et al., 2001; O'Connor et al., 1998). As other research indicates a number of variables such as financial stress, family type, quality of marital relationship and family/neighbourhood violence impacts on attachment (Op.Cit). These factors need controlling for in studies to ascertain effects.

Finally the relevance of genetics and environmental factors combined will need closer scrutiny (Deater-Deckard, 2000; Glaser, 2000; O'Connor et al., 1998). As Glaser (2000) highlights, *"Developmental Psychology and the study of behaviour and emotion have tended to be considered in parallel to the study of neurobiological processes"* (p.97). This is true in relation to Nurture Group research, with the potential role of attachment on neurobiology ignored. In future these issues will need considering together, with the role of genetics placed within the context of the attachment relationship in Nurture Group interventions (O'Connor et al., 1998). In

practice, the combination of biological and environmental factors impact could be considered using an interactive factors framework such as causal modelling (Morton & Frith, 1995).

International Level – Long Term Action to Explore Future Research Directions

There exists a range of research possibilities within the Nurture Group domain to develop attachment theory further and consider alternative interventions. In terms of attachment there is an *"ongoing need to incorporate research findings from social, genetic and developmental psychology"* (O'Connor et al., 1998) (p. 979). This is certainly the case in Nurture Group research where a combination of related disciplines may help illuminate the mechanisms at work. As Deater-Deckard (2000) argues, future research will need to be *"longitudinal and will examine environmental and gene-environment process"* (p. 482). The role of Nurture Groups will need to be considered alongside such research with later potential for change ascertained.

Nurture Group research will need to carefully evaluate the role of attachment or other mechanisms as the critical variables in the intervention. Current literature has made little attempt to uncover the mechanisms at work underlying Nurture Group provision. This could be overcome by longitudinal designs with carefully matched controls. Other work has indicated the role of attachment in buffering infants' response to stress and as a protective factor for abused children (Glaser, 2000). Nurture Group research will have to

ascertain the critical component of the intervention and whether this pertains to attachment. Furthermore, as research in resilience indicates, the extent to which Nurture Groups are a mediating variable (Falkov, 1998) or interactive variable (Fonaghy et al., 1994) will need to be determined.

Alternative interventions, which aim to remedy early attachment difficulties, could be compared alongside Nurture Groups in future research. As Belsky and Fearon (2002) highlight, "*few studies especially of early attachment, have tested interactive propositions regarding the moderating role of later experiences on the predictive power of early development*" (p. 380). Nurture Groups could be compared with matched samples to other interventions that aim to ameliorate early attachment issues. Alternative interventions focussing on the parents as the key agents of change could be examined. The Circle of Security Project (Marvin et al., 2002) is one such example to act as a comparison. The intervention is a "*20 week, group based, parent education and psychotherapy intervention designed to shift patterns of attachment-care giving interaction*" (Marvin et al., 2002) (p. 107). Another intervention for comparison could be the Sunderland Infant Programme which aims to "*support parents in helping their babies to develop a secure attachment*" using a multi-professional approach (Maher, 2002) (p. 18). These earlier interventions could be compared longitudinally with matched samples to Nurture Groups. This would be informative as to the locus of the intervention and the impact of time, duration and optimum age. As Glaser (2000) details *"Future prospective research will need to continue to test the hypothesis that a greater and earlier direct investment in children leads to enduring improvement in their functioning"* (p. 110). In practice cost

comparisons in relation to Nurture groups and alternatives could be examined to ascertain the effectiveness of early intervention over time.

Future avenues for development could also investigate the impact of interventions when parents are the focus for change. This could explore the impact of parenting skills (Iwaniec & Herbert, 1999) or concepts such as emotional availability (Bretherton, 2000). In these cases, the most appropriate intervention vehicle, be it home or school could be determined. In practice such alternatives could be explored to gauge effectiveness at an initial qualitative level. Furthermore the development of a parent-training program for parents who have children in Nurture Groups could be used to complement the intervention. EP's would be ideally placed to put this into practice alongside existing Nurture Groups.

Concluding Comments

This paper demonstrated the relevance and importance of Nurture Groups at a local and national level. The literature review parallels such interest with a range of research examining Nurture Group effectiveness. Despite the existence of some positive effects, however, considerable research is still required to ascertain the impact on children and schools. The need to address present shortfalls between the domains of research and practice can be achieved by short, medium and long-term action. In particular, Educational Psychology as a discipline can provide the relevant scientific rigour warranted to reveal the true evidence of effectiveness for Nurture Groups.

References

Ainsworth, M.D.S. (1967). *Infancy in Uganda: infant care and the growth of love.* Baltimore, M.D.: Johns Hopkins Press.

Ainsworth, M.D.S., Blehar, M.C., Waters, E. and Wall S. (1978). *Patterns of attachment: a Psychological Study of Strange Situations.* Hillsdale, NJ: Erlbaum.

Balbernie, R. (1999). Infant Mental Health. *Young Minds,* 39, March/April, 1999.

Bartholomew, K., and Moretti M. (2002). The dynamics of measuring attachment. *Attachment and Human Development,* 4(2), 162-165.

Belsky, J. and Fearon, R.M.P. (2002). Early Attachment Security, subsequent maternal sensitivity, and later child development: Does continuity in development depend upon continuity of care giving? *Attachment and Human Development,* 4(3), 361-387.

Bennathan, M. (1997). Effective Intervention in Primary Schools: What Nurture Groups Achieve. *Emotional and Behavioural Difficulties,* 2(3), 23-29.

Bennathan, M. and Boxall, M. (1998). *The Boxall Profile: a guide to the education of pupils with emotional and behavioural difficulties.* Maidstone Kent: AWCEBD.

Bishop, A. and Swain, J. (2000a). 'The Bread, The Jam and Some Coffee in the morning': Perceptions of a Nurture Group. *Emotional and Behavioural Difficulties*, 5(3), 18-24.

Bishop, A. and Swain, J. (2000b). Early years education and children with behavioural and emotional difficulties: Nurturing parental involvement? *Emotional and Behavioural Difficulties*, 5(4), 26-31.

Bowlby, J. (1951). *Maternal Care and Mental Health*. WHO Monograph Series, No.2. Geneva: World Health Organisation.

Bowlby, J. (1953). *Child Care and the Growth of Love.* London: Pelican.

Bowlby, J. (1969). *Attachment and loss: Vol.1 Attachment*. London: Hogarth Press.

Bowlby, J. (1980). *Attachment and Loss, Vol.3. Loss, sadness and depression*. London: Hogarth Press.

Bretherton, I. (2000). Emotional Availability: an attachment perspective. *Attachment and Human Development*, 2(2), 233-241.

CSIE (2000). *Index for Inclusion.* Bristol: Centre for Studies on Inclusive Education.

Connor, M.J. (2001). *The nature of nurture groups and evidence for effectiveness.* Surrey Educational Psychology Service, November 2001.

Cooper, P. (2003). Including students with social, emotional and behavioural difficulties in mainstream secondary school. *Emotional and Behavioural Difficulties*, 8(1), 5-6.

Cooper, P. and Lovey, J. (1999). Early intervention in emotional and behavioural difficulties: the role of Nurture Groups. *European Journal of Special Needs Education*, 14(2), 122-131.

Cooper, P. Arnold, R., and Boyd, E. (2001). The effectiveness of Nurture Groups: preliminary research findings. *British Journal of Special Education*, 28(4), 160-165.

Deater-Deckard, K. (2000). Parenting and Child Behaviour Adjustment to Early Childhood: A Quantitative Genetic Approach to Studying Family Processes. *Child Development,* 71(2), 468-484.

Department for Education and Employment (1997). *Excellence for all children: Meeting Special Educational Needs.* London: HMSO.

Department for Education and Employment (1998). *Meeting Special Educational Needs: A programme for action.* London: DfEE.

Department for Education and Employment (1999a). *Social Inclusion: Pupil Support*. Circular 10/99. London: DfEE.

Department for Education and Employment (1999b). *Social Inclusion: The LEA Role in Pupil Support* Circular 10/99. London: DfEE.

Department for Education and Employment (2000). *Educational Psychology Services (England): Current Role, Good Practice and Future Directions.* The Research Report. London: DfEE.

Doyle, R. (2001). Using a readiness scale for reintegrating pupils with social, emotional and behavioural difficulties from a Nurture Group into their mainstream classroom – a pilot study. *British Journal of Special Education,* 28(3), 126-131.

Falkov, A. (1998). *Crossing Bridges. Training Resources for working with mentally ill parents and their children. Reader – for managers, practitioners and trainers.* Pavilion Publishing, Brighton.

Fonaghy, P., Steele, M., Steele, W., Higgett, A. and Target, M. (1994). The Emmanuel Miller Memorial Lecture, 1992: The theory and practice of resilience. *Journal of Child Psychology and Psychiatry*, 35(2), 231-257.

Glaser, D. (2000). Child Abuse and Neglect and the Brain – A Review. *Journal of Child Psychology and Psychiatry*, 41(1), 97-116.

Geddes, H. (2003). *Attachment Theory*. A presentation at Surrey Educational Psychology Service Conference, 20th March, 2003.

George, C., Kaplan, N., and Main, M. (1985). *The Adult Attachment Interview*. Department of Psychology, University of California, Berkeley.

Goldberg, S., Muir, R. and Kerr, J. (eds) (1995). *Attachment Theory: Social, Developmental and Clinical Perspectives* – Hillsdale, NJ: The Analytic Press.

Green, J., Stanley, C., Smith, V., and Goldwyn, R. (2000). A new method of evaluating attachment representations in young school-age children: The Manchester Child Attachment Story Task. *Attachment and Human Development,* 2(1), 48-70.

Greenberg, M.T. (1999). Attachment and Psychopathology in Childhood. In Cassidy, J., and Shaver, P.R. (Eds), *Handbook of Attachment; Theory, Research and Clininical Applications* (pp 469-496.) New York; Guilford Press.

Greenfield, S. (1997). *The Human Brain – a guided tour*. London: Phoenix.

Greenhalgh, T. (2001). *How to Read a Paper: The basis of evidence based medicine.* London: BMJ Books.

Head, G., Kane, J. and Cogan, N. (2003). Behaviour support in secondary schools. *Emotional and Behavioural Difficulties*. 8(1), 33-42.

Hayden, C. (1997). Exclusion from primary school: children "in need" and children with "special educational need". *Emotional and Behavioural Difficulties* 2(3), 36-44.

http://www.attachmentnetwork.org (Attachment Network).

http://www.Nurturegroups.org (Nurture Group Network).

http://www.tandf.co.uk/journal/titles/14614734.hton (Attachment and Human Development Network).

Iszatt, J. and Wasilewska, T. (1997). Nurture Groups: An early intervention model enabling vulnerable children with emotional and behavioural difficulties to integrate successfully into school. *Educational and Child Psychology*, 14(3), 63-70.

Iwaniec, D. and Herbert, M. (1999). Multi-dimensional approach to helping emotionally abused and neglected children and abusive parents. *Children and Society,* 13, 365-379.

Kobak, R., Little, M., Race, E. and Acosta, M.C. (2001). Attachment disruptions in seriously emotionally disturbed children: Implications for treatment. *Attachment and Human Development*, 3(3), 243-258.

Lloyd, G., Stead, J. and Kendrick, A. (2003). Joined-up approaches to prevent school exclusion. *Emotional and Behavioural Difficulties*, 8(1), 77-91.

Lucas, S. (1999). The Nurturing School: The Impact of Nurture Group principles and practice on the Whole School. *Emotional and Behavioural Difficulties*, 4(3), 14-19.

Mackay, T. (2002). Discussion Paper – The future of Educational Psychology. *Educational Psychology in Practice*, 18(3), 245-253.

Maher, T. (2002). Talking Babies: Securing early attachment. *Young Minds*, 60, 18-19.

Marvin, R., Cooper, G., Hoffman, K. and Powell, B. (2002). The Circle of Security Project: Attachment-based intervention with care-giver/pre-school child dyads. *Attachment and Human Development*, 4(1), 107-124.

Morton, J. and Frith, U. (1995). Causal modelling: a structured approach to developmental psychopathology. In Cicchetti, D. and Cohen, D.J. (eds) *Manual of Developmental Psychopathology* (pp. 357-90). NY Psychological Assessment of Dyslexia: Wiley

O'Connor, T. and Colwell, T. (2002). The effectiveness and rationale of the "nurture group" approach to helping children with emotional and behavioural difficulties remain within mainstream education. *British Journal of Special Education,* 29(2), 96-100.

O'Connor, T.G., Deater-Deckard, K., Fulker, D., Rutter, M. and Plomin, R. (1998). Genotype – Environment Correlations in Late Childhood and Early Adolescent: Antisocial Behavioural Problems and Coercive Parenting. *Developmental Psychology*, 34(5), 970-981.

Ramchandani, P., Joughlin, C. and Zwi, M. (2001). Evidence-based child and adolescent mental health services: oxymoron or brace new dawn. *Child Psychology and Psychiatry Review,* 6, 59-64.

Robertson, I. (1999). *Mind Sculpture.* London: Bantam Press.

Rudestam, K.E. and Newton, R.R. (1992). *Surviving your Dissertation.* Newbury Park, California: Sage.

Rutter, M. (1995). Clinical Implications of Attachment Concepts: Retrospect and Prospect. *Journal of Child Psychology and Psychiatry,* 36(4), 549-571.

Shearman, S. (2003). What is the reality of "inclusion" for children with emotional and behavioural difficulties in the primary classroom? *Emotional and Behavioural Difficulties*, 8(1), 53-76.

Steele, H. (2002). State of the Art: Attachment Theory. *The Psychologist*, 15(10), 518-522.

Svanberg, P.O.G. (1998). Attachment, resilience and prevention. *Journal of Mental Health*, 7(6), 543-578.

Trevarthen, C. and Aitken, K.J. (2001). Infant Intersubjectivity: Research, Theory and Clinical Applications. *Journal of Child Psychology and Psychiatry*, 42(1), 3-48.

Turner, M (2003). *Recent Research in Attachment Theory – Implications for Teaching and Learning*. A presentation at the National Conference for Educational Psychologists enrolled in Doctoral Programmes, 4th July 2003.

Professional Practice Assignment 3: Supervision for Educational Psychologists in Training (EPiTs): Preparing for the Future

Abstract

EPiT supervision is an area of considerable and crucial importance in the profession of Educational Psychology being the focus of renewed interest given the imminence of extended training (Malcolm, 2003; Turner, 2003). This paper examines the present research literature to ascertain prevalent practice in LEA contexts. Such evaluation reveals that, although, relevance is recognised considerable variation exists in EPiT supervision provided. Indeed, research is rife with a number of unresolved issues and unanswered practice questions. Recommendations for future research and practice are highlighted to pave the way forward for the development of quality EPiT supervision. As Nolan (1999) argues "the future of supervision is in our hands" (p.106).

Section 1: Aims and Scope of Assignment

Supervision is increasingly acknowledged as a fundamental tool essential for the continuing professional development of Educational Psychology (Carrington, 2004; Lunt, 1993). This has been evidenced in both research (Hamilton-Farrell, 1993; Jennings, 1995; Lunt & Sayeed, 1995; Nash, 1999; Nolan, 1999; Pomerantz & Lunt, 1993) and by professional bodies (BPS/DECP, 2000; 2002). The reasons for this pertain to the intrinsic value of the supervisory process. Indeed, supervision has a variety of functions for practitioners, from support (Dowling & Osborne, 1994; Hawkins & Shohet,

2002) to providing opportunities for active reflection (Jennings, 1995; Vecchio & Boatwright, 2002). In this manner supervision is becoming recognised as "*an important component in ensuring high quality professional practice in Educational Psychology*" (Nolan, 1999) (p. 98). Benefits, likewise, exist at an organisational level within the research literature (Arnold & Johnson, 1997; Joiner et al., 2004; Kleinman et al., 2001; Orpen, 1997).

Renewed interest in supervision has been apparent from professional bodies such as the DECP and BPS with the production of Quality Standards for Educational Psychology Services (October, 2000) and Professional Practice Guidelines (January, 2002). Moreover, the imminence of extended training will necessitate substantial change in the provision of supervision from both LEAs and HEIs (Malcolm, 2003; The Tavistock & Portman NHS Trust, 2003; Turner, 2003). All these factors illustrate the growing relevance of supervision in the profession. This piece aims to explore one important aspect of supervision, namely that of EPITs. Such a specific focus was adopted as it represents a need for development within the discipline (Malcolm, 2003; Turner, 2003). Furthermore, the Association of Educational Psychologists has raised concerns in this domain (AEP Circular, 2004a). A more extensive exploration of supervision in general did not occur as it was considered too large to examine in the context of this assignment.

Definitions

As Pomerantz (1993a) argues "*supervision is not a unitary social activity with tightly prescribed traditions, boundaries and structures. It varies considerably and this affects our understanding of it*" (p. 16). Indeed the concept is ill-defined as Nolan (1999) describes "*supervision is a complex multi-functional concept*" (p. 98). For the purposes of this assignment, the following definition will be used as proposed by the DECP "*a process of examining one's own work and issues arising from it, at a professional and personal level*" (cited in Osborne, 1993) (p. 27). Furthermore, the joint problem-solving aspect of supervision will be encapsulated by Nolan's (1999) definition, "*supervision is an interactive process with responsibility often being shared between supervisor and supervisee*" (p. 100). Finally, Hawkins and Shohet's (2002) three main purposes of supervision will be referred to. These relate to functions from educative in the development of skills, supportive in dealing with stress and managerial in relation to quality assurance.

Overview

It is clear from recommendations made within the profession (BPS/DECP, 2002; DfEE, 2000) and given proposed changes in training (Malcolm, 2003) that alteration to the fieldwork supervision of EPiTs will be required. This assignment intends to examine present research and practice to outline future directions. To begin an overview of the context of supervision of EPiTs at a local and national level will be illustrated. A critique of the existing research will then follow, which in turn will inform discussion into implications for theory and practice in the wider context.

Section 2: Practice and Context

The impetus for investigating the fieldwork supervision of EPiTs exists at a number of levels from both local to national (Lunt, 1993; Nolan, 1999; Webster et al., 2000).

Local Context

In a large county Educational Psychology Service a small working party of EPs aimed to evaluate the current supervisory processes being utilised. The rationale for investigation was outlined in a consultation document sent out to all practitioners in December 2003. These principally related to the changing requirements of chartered status alongside the imminence of extended training (Malcolm, 2003). Following this initial phase, a needs analysis (Binsted, 1980) was conducted with individuals. Further information was gathered through consultation with area teams. In this manner it was hoped individual EP views would be considered in context.

The key findings illustrate a need in the local context to develop current practice alongside the knowledge base with regard to EPiT supervision. Data indicated that it tended to be distributed according to areas on an ad hoc basis. Individuals would use the models delineated by the training institution concerned. In this manner practice was isolated and exposed to the potential mismatch of espoused university theory contrasted to the theory in use in the locality (Argyris & Schon, 1978). Furthermore, the competencies of practitioner's knowledge were varied in terms of both the skills and confidence in supervising EPiTs (Falender et al., 2004). This

included the models and methods of supervision to use, alongside the processes involved.

In addition, feedback highlighted the need for time allocation to both give and receive supervision. In a similar manner it was considered that the *"notion of the importance of supervision is not properly embedded in the culture"* (Surrey Children's Services, 2004) (p.2). This was apparent given no documents on supervision for EPiTs or EPs existed. This may have an effect on EPiT supervision for as Pomerantz (1993b) argues *"undoubtedly, if more practising EPs had access to regular, high quality supervision it would benefit trainee EPs on placement"* (p. 34). This is of growing importance given *"the probable development of 3 year training which will increase the responsibility of LEA supervisors on practice"* (The Tavistock & Portman NHS Trust, 2003) (p.1). In this manner, at a local level, both the knowledge and practice of EPiT supervision requires development. As Gersch (2001) highlights such supervisory structures will benefit not only EPiTs, but all practitioners in supporting continuing professional development through active and critical reflection.

At a local level this work offered a consultative evaluation of supervision within the service at both a practitioner and team perspective. It usefulness is limited, however, as procedures and measures used were not reliable or valid statistically. Similarly the lack of any pilot study in the experimental design means a range of extraneous variables were unaccounted for. Likewise, the systematic bias on behalf of the working group may have

meant reflexivity was a confounding variable (Willig, 2001). Finally, the absence of a clear qualitative technique for both data collection and interpretation restricts credibility (Potter & Wetherell, 1987). Despite these limitations the consultation process aimed to provide an exploratory insight into local practice of EPiT supervision.

National Context

On a national level the rationale for this assignment originates from two levels, principally from the development of Educational Psychologists in Training to the evolution of the profession as a whole.

Development of Educational Psychologists in Training

Supervision is recognised as crucial in the support and development of trainees on the pathway to professional practice (Carrington, 2004: Nash, 1999; Nolan, 1999; Scaife, 2001). As Lunt (1993) cites *"supervising is immortality: it leads to the next generation of practitioners"* (Houston, 1990 in Lunt, 1993a) (p.3). Supervision is regarded as the main mechanism for supporting trainees and is of *"importance to the quality and outcome of training"* (Scaife, 1993) (p.61). In this manner supervision supports the trainee on a number of levels from professional development in the educative sense to the supportive in assisting with stress and challenge (Hawkins & Shohet, 2002). With regards to professional development, supervision allows for *"the phenomenon of the theory practice dilemma"* faced by many EPiTs to be addressed (Bromme & Tillema, 1995) (p. 261). As Hawkins and Shohet (2002) detail supervision *"can give us a chance to stand back and reflect"* (p.

3). It is such guided reflection that can provide EPiTs with the professional artistry needed to blend theory and practice seamlessly together (Lunt, 1993b). As Carrington (2004) highlights, this supervision is crucial in the move from *"novice to professional"* (Eraut, 1994 in Carrington, 2004).

Supervision is similarly supportive for EPiTs in managing the multifaceted nature of the EP role and its associated stresses. As Kuk and Leyden (1993) highlight *"the educational psychologist has to balance the complexities of working with children, parents, teachers and systems at a range of levels and integrating the respective practical, theoretical, emotional and ethical issues"* (p. 51). Supervision is a means by which EPiTs can review how to manage these, for as Hawkins and Shohet (2002) illustrate it is a *"central form of support"* (p. 23). Indeed Lunt and Sayeed (1995) detail the importance of supervision in enabling practitioners to *"carry out their professional role and survive the demands of the job"* (p. 26). As Buchanan and Huczynski (1997) point out, once individual practitioner needs are met they will be able to provide for clients accordingly. In this manner supervision fulfils a range of key functions for EPiTs as epitomised by the following quote *"Supervision may provide the opportunity for professional support, the wider perspective, the processing and articulation of professional dilemmas and the space to develop 'reflective practice' which may provide a more robust foundation for our professional work"* (Lunt, 1993) (p. 3).

Despite the critical nature of supervision for EPiTs and the profession, a number of constraints exist on current practice. As Scaife (1993; 2001)

highlights the frameworks for such a process tend to be based on teaching models from education as opposed to those from clinical or occupational practice. Indeed, a range of work in the area suggests disparities and inconsistencies of supervision in practice (Hamilton-Farrell, 1993; Lunt, 1993; Nolan, 1999; Osborne et al., 1990). It appears a range of alternative practices exist with high variability in experiences from EPiTs. As Powell et al (1990) indicate, although *"we have taken a long step away from the notion of fieldwork supervision as the apprenticeship component of training....there is as yet no model established which is specifically designed for our own profession"* (cited in Hamilton-Farrell, 1993). As such, this represents a significant area to develop for the future support of EPiTs.

Development of Educational Psychology as a Profession

Supervision is recognised as of increasing importance in order to evolve as a profession in accordance with other strands of psychology. As Lunt (1993b) points out "*unlike other professions such as social work, educational psychology has not had a history or culture of supervision as part of its professional work*" (p. 4). A number of changes, however, have been made in the last four years to address this imbalance and align the profession with occupational counterparts. The DfEE Report (July 2000) alongside quality standards and professional practice guidelines outlined by the BPS were instrumental in this (BPS/DECP, 2000; 2002). As one institution indicates "*The BPS Quality Standards for Educational Psychology Services make clear the increasing importance that is attached to supervision*", (The Tavistock &

Portman NHS Trust, 2003) (p. 1). Supervision is now considered a universal requirement for all members of the profession.

An additional impetus for promoting the salience of supervision in the profession is the development of extended training. Such changes will necessitate a far greater degree of supervision over a lengthy time period for EPiTs professional induction (Malcolm, 2003; Turner, 2003). In particular LEAs "*would be required to participate in the professional supervision and support of EPiTs*" over a prolonged period (Malcolm, 2003). Supervision will need to be of a high standard, at least 30 minutes a day, and will become a mandatory requirement. Such supervision is proposed to occur though field/practice tutors who will be concerned with the personal and professional development of EPiTs within placement contexts. Such a move will need change at an organisational LEA level for Educational Psychology Services (Turner, 2003). Such changes have generated anxieties from the Association of Educational Psychologists, in particular *"the logistics of placements including arrangements for supervision*" (AEP Circular, 2004b). These issues highlight the current and controversial concerns surrounding EPiT supervision in the profession.

Considerable agreement exists within the profession as to the importance of supervision from individual practitioners locally, to the DECP, BPS and AEP perspectives nationally. Controversy, however, remains as to the espoused theory and that in use, with practice neither universal nor established (Hamilton-Farrell, 1993; Pomerantz, 1993a; 1993b). Moreover, in terms of

research there is still a great need for further examination (Hamilton-Farrell, 1999; Nash, 1999; Nolan, 1999; Pomerantz & Lunt, 1993). Indeed this is further necessitated given the stark comparison with fellow professionals in the realm of supervision (Lunt, 1993b; Lunt & Sayeed, 1995; Pomerantz, 1993a; 1993b). Moreover, the onset of extended training heralds a need for developments in this domain. This assignment seeks to illustrate the potential of EPiT fieldwork supervision in particular, for the future advancement of the profession.

Section 3: Psychological Theory and Research

The literature search used bibliographic databases of PsychInfo and ERIC as recommended by Ramchandi et al (2001). The key search terms used were "*Supervision*", "*Supervisory Frameworks*" and "*EPiT Supervision*". Manual searches from recent journals occurred in conjunction with ancestral searches.

The review aims to critique the research literature surrounding the practice of EPiT supervision in LEA Educational Psychology Services in the United Kingdom. The research to be examined was, therefore, selected from UK samples and included evidence from EPiTs, newly qualified and established practitioners in the profession. The review uses the existing research to date as reported in the field (Carrington, 2004; Lunt, 1993; Nolan, 1999; Webster et al., 2000). The critique is divided into 5 sections using the critical evaluation checklist from Rudestam and Newton (1982). Initially the theoretical frameworks and conceptualisations in the literature are explored.

Following on from this the research design is critiqued alongside the results and discussion. A final section details the major themes and controversies.

1. Conceptualisation

Studies in the EPiT supervision literature identified the major area of investigation to be the issues, processes and practices of supervision in relation to the LEA context. The National Survey by the DECP Training Committee and Course Tutors Group aimed to *"investigate the whole process of fieldwork supervision"* from the perspectives of EPiTs, Supervisors and Course Tutors (Osborne et al., 1990) (p.37). Similar studies from Hamilton Farrell (1993) and Thompson (1993) focused in on EPiTs in a similar vein by surveying experiences from one trainee cohort. Work from Lunt and Sayeed (1995) and Webster et al (2000) extended such research by examining perceptions of newly qualified educational psychologists concerning supervision and induction, during the transition period. Finally other studies provided an individual and reflective insight into practitioners who had supervised EPiTs including Carrington (2004); Ferguson (1990) and Nash (1999). All studies identified an issue for investigation accurately in the context of Educational Psychology. In particular the need for exploration was highlighted given supervision provides the *"powerful blend of theory and practice"* to enable trainees to *"make the transition from one profession to another"* (Osborne et al., 1990) (p.37). Furthermore all studies used qualitative research methodologies to examine the processes involved. In summary much cumulative development is apparent in the research with

later studies building on earlier ones to offer insight into the practice of supervision.

Overall the process of supervision was made explicit alongside its role and purpose in the profession (Carrington, 2004; DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Webster et al., 2000). A number of studies delineated a range of models in relation to supervision from the three part model of Hawkins and Shohet (1990) in Hamilton-Farrell (1993), to models of application as typified in the DECP (1987) study and in Nash (1999). In terms of theoretical underpinnings, however, evidence varied with some studies omitting the origins of supervisory concepts (Ferguson, 1990; Nash, 1999; Thompson, 1993). This indicates a restricted theoretical research base which impacts on credibility. Despite this, the DECP study in particular, defined supervision clearly and even made comparisons with other professions to provide an impetus for investigation (DECP, 1987; Osborne et al., 1990; Pomerantz, 1990).

2. Theoretical Framework and Hypotheses

With regards to theoretical frameworks the research varied according to the models adopted. A number of pieces utilized definitions and models from the helping professions namely Hawkins and Shohet's (1990) three function framework differentiating between managerial, educative and supportive supervision (cited in Carrington, 2004; Hamilton-Farrell, 1993; Nash, 1999). Other studies focused on different modes of supervision such as the DECP

(1987) and Ferguson (1990). These studies did not, however, evidence on what these distinctions were founded. Moreover, as Carrington (2004) highlights, few studies focus on supervision as a reciprocal learning process from the theory base.

The research from both the DECP (1987) and Webster et al (2000) had clearly based research questions generated from an overview of the literature. These linked to the design of measures and the method in a clear and coherent manner. Similarly both pieces operationalised hypotheses explicitly. Other studies, however, whilst able to provide a general intention of investigation were unable to specify hypotheses in a testable manner. The questionnaire surveys of Hamilton-Farrell (1993); Lunt and Sayeed (1995) and Thompson (1993) lacked hypotheses, making interpretation difficult. Likewise, the relationships amongst variables were unclear and unaccounted for. Lunt and Sayeed (1995) did, however, link research to earlier studies and Hamilton-Farrell (1993) did include some research questions. The individual accounts also had an absence of research questions and hypotheses. These pieces rather provided an illuminative individual description of supervisory experience. In this sense the majority of studies lacked the necessary scientific rigour to delineate hypotheses and variables in explicit form, which undermines credibility.

3. Research and Design

The majority of research in the area used semi-structured questionnaires with the addition of focus groups or interviews to examine EPiT supervision (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Thompson, 1993; Webster et al., 2000). These techniques were appropriate to the field of research being addressed and justified given the sample size involved. Few studies, however, provided a rationale to explain the use of a questionnaire approach or the theoretical foundations underpinning the research methodology. In particular, no discussion occurred surrounding the potential use of a quantitative techniques or additional qualitative methods. This represents a narrow perspective from which to examine supervision. Furthermore, the manner in which researcher bias would be minimised in the construction, collection and interpretation of data was not addressed. The individual pieces based on practitioner supervisors were similarly limited, with no explanation of research design making replication impossible (Carrington, 2004; Ferguson, 1990; Nash, 1999). The failure in current research remains in the inadequacy in providing sufficient information on the methodology adopted. In the future a clear procedure and rationale for design would ameliorate these difficulties.

In terms of extraneous variables, no study adequately controlled for the range of confounding factors. Some pieces such as the DECP (1987), Webster et al (2000) and Hamilton-Farrell (1993) discussed variables involved, but as in all other work, failed to reasonably operationalise these in

research design. Moreover, confounding variables were not accounted for by any piece in terms of researcher bias evident in the design and interpretation of data. As Potter and Wetherall (1987) argue *"people are using their language to construct versions of the social world"* (p.33). Questionnaires were not blind and no research evidenced any safeguards to protect against subjectivity in the qualitative methodology, which may have skewed results and interpretations. Future work would need to address researcher reflexivity, for example making data collection and analysis explicit (Yin, 1994 in Sherrard, 1997) and raw data available (Erlandson et al., 1993 in Sherrard, 1997).

Details of measurement devices used in the studies discussed is variable. In the DECP (1987) and Webster et al's (2000) research, information is provided as to questionnaire construction, linking the design to earlier work. In contrast other questionnaire surveys such as Hamilton-Farrell (1993); Lunt and Sayeed (1995) and Thompson (1993) gave no detail on the selection or construction of the instrument. Indeed in all studies technical detail regarding validity and reliability statistics with regard to questionnaires is omitted. In this sense measures were not objectively validated with ethical considerations ignored as to the diversity of the population on whom the measure had been standardised. No assessment of the measures completeness, quality or strength was evidenced by external validation. In future specific statements relating to the construction of questionnaires alongside the distribution and counter balancing of statements would be required. Likewise the impact of the researcher in both the construction and

interpretation of questionnaires would be needed. The individual studies also had no technical detail on the measures adopted (Carrington, 2004; Ferguson, 1990; Nash, 1999). At present research has been limited by the quality and range of measurement devices utilised.

With regard to the sample the questionnaire surveys used participant populations appropriate and relevant to the research area being studied. The DECP Study used EPiTs, Supervisors and Course Tutors to sample all perspectives. The work of Hamilton-Farrell (1993); Lunt and Sayeed (1995); Thompson (1993) and Webster et al (2000) all targeted EPiTs. Such research is limited, however, by a number of factors. Firstly all the studies represent findings from a one year cohort of practitioners alone, that makes generalisation of results questionable. In addition the process of participants' recruitment lacks detail with inclusion/exclusion statistics omitted. Moreover, the demographic characteristics of the sample are not reported on, which makes it difficult to ascertain how representative the cohort was. Most fundamental is the fact that response rates varied from 47% to 75% which could mean only a skewed view of supervision was obtained (Belsky & Fearon, 2002). The sample size was also not calculated. In future these factors would need addressing to enhance the credibility of study in the area. The individual practitioner perspectives are limited in terms of sample size as they represent only one viewpoint (Carrington, 2004; Ferguson, 1990; Nash, 1999).

4. Results and Discussion

Studies investigating EPiT supervision varied in the manner in which results were reported. The DECP (1987) study organised findings into themes underpinning questionnaire responses in keeping with the qualitative data collected. Other studies used percentages of responses to questionnaire items with individual quotes to illuminate meaning (Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Thompson, 1993). Overall little detail was evident in the research as to the qualitative techniques selected and rationale for use. Moreover, the processes of data treatment in terms of transcription and coding are omitted. As Willig (2001) indicates *"data must be coded, summarised, categorised or otherwise "reduced" at the point of collection"* (p.16). This translation and data analysis is lacking with the process of developing codes and themes unspecified. The absence of any clear technique in analysis and interpretation reduces both credibility and replicability. Similarly issues of experimenter reflexivity are not addressed restricting validity and reliability. These limitations are also apparent in the practitioner perspectives of Carrington (2004); Ferguson (1990) and Nash (1999). As a consequence the research inadequately manages and accounts for confounding variables in the data analysis. An additional flaw is the over-generalisation of results based on limited response rates.

With regard to conclusions drawn, a number of studies have discussions, which are consistent with results. The pieces from the DECP (1987) and Webster et al (2000) make conclusions, which directly relate to data collected

and link back to earlier hypotheses. Likewise, alternative conclusions consistent with the data are examined. Other studies, whilst able to generate conclusions connected to results, fail to acknowledge alternative interpretations (Carrington, 2004; Ferguson, 1990; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Thompson, 1993). This may lead to a skewed interpretation with causality and inferences made by the researcher with little evidence from the data. Limitations in the research are identified by the DECP (1987) in relation to the use of a questionnaire and by Hamilton-Farrell (1993) in terms of the sample size and representativeness. Few other studies discuss or refer to any restrictions of the research and no paper recognises the role of confounding variables. Theoretical and practical implications are delineated by the work of the DECP (1987); Thompson (1993) and Webster et al (2000) whilst other studies report on the practice issues alone (Carrington, 2004; Ferguson, 1990; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Nash, 1999).

5. Summary of Psychological Literature

All studies provided a useful contribution into the research of EPiT supervision. The DECP (1987) study is the most methodological sound and makes recommendations for both theory and practice linked to fellow professionals. It like all other pieces, however, was limited by its lack of detail pertaining to qualitative techniques. The studies of Hamilton-Farrell (1993) and Thompson (1993) offered a useful insight into EPiTs specifically as did Lunt and Sayeed (1995) and Webster et al (2000) into newly

qualifieds. Despite this, the adequacy of the studies were restricted by the sample size and response rates. The research of Carrington (2004), Ferguson (1990) and Nash (1999) demonstrated an illuminating perspective into supervision from that of LEA supervisors. Considerable issues with the case study level of analysis exist, however, which limit credibility. Research in this domain needs to be developed with respect to experimental design in particular.

With regard to the major themes evident in the psychological literature a number of controversies were apparent in relation to the variability of EPiT supervision. Many studies noted that EPiTs had different degrees of access to supervision (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995). As Powell et al (1990) pointed out there was *"concern about the variability of supervision received by trainees"* (p.44). Indeed time appeared to be a fundamental factor in this with *"wide variation in the amount of supervision (formal and informal) received by trainees on placement"* (Lunt, 1993b) (p.9). The competing pressures on time meant supervision varied considerably (DECP, 1987; Hamilton-Farrell, 1993; Lunt, 1993b; Lunt & Sayeed, 1995; Webster et al., 2000). As Osborne et al (1990) indicated 21% of EPiTs spent less than 10% of time in supervision with a clear *"variation in the amount of time allocated to supervision of trainees on placement"* (p.39). Furthermore, research indicated the wide variety of types of EPiT supervision available (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Thompson, 1993). As Osborne et al (1990) noted there was *"wide variation in practice, and disturbingly, a number of trainees never having had the chance to*

observe a professional role model. Nor had they received critical feedback based on direct observation" (p.40). All research evidenced the disparity in terms of type, time and access with regard to EPiT supervision in LEA contexts.

An additional major theme consistent across research was the lack of training for EPiT supervision (DECP, 1987; Ferguson, 1990; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995). As Carrington (2004) highlights,*" this has been an overlooked area in the past"* (p.40). The research indicated this absence of training from perspectives of both EPiTs (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995) and supervisors' alike (Carrington, 2004; DECP, 1987; Ferguson, 1990). It could be argued that the variability in time, type and access to supervision alluded to earlier could be a product of this. As Powell et al (1990) argues, *"such variability must arise, surely, from the lack of a common training in Supervision"* (p.44). Indeed this factor has been singled out as one of the most crucial in creating qualitative variation in EPiT supervision across placements.

A key area identified in all studies was the recognition of complex ethical issues, which may arise in EPiT supervision. In particular the potential power imbalance between EPiTs and Supervisors was highlighted (DECP, 1987; Hamilton-Farrell, 1993; Nash 1999). As Nash (1999) indicated, such a *"power imbalance can lead to difficulties in building effective working relationships"* (p.113). These inequalities could be listed from differing attitudes (Nolan, 1999) to *"assumptions of supervisors which related to*

gender, race or age" (Hamilton-Farrell, 1993) (p.77). As Carrington (2004) indicates *"the tensions in the supervisory relationship are widely acknowledged"* (p.39). Such issues consistently arose in the research surrounding EPiT supervision. As Powell et al (1990) point out these factors need careful consideration given the potential impact they may have on *"the development of the trainee's confidence and self-image as a professional psychologist"* (p.49).

A number of key practice questions were identified as unanswered in the research literature. The first of these relates to the blurred distinction concerning the purpose of EPiT supervision in the LEA context in terms of the dual role of assessment and support. As Powell et al (1990) highlight *"the greatest area of potential conflict derives from the supervisors split role in juggling responsibility for the client with responsibility for the trainee's learning, as well as acting as assessor of it"* (p.51). Indeed this combined role is detailed in a range of research as controversial (Carrington, 2004; Ferguson, 1990; Kuk & Leyden, 1993; Lucas, 1989). In particular Pomerantz, (1990) discusses how *"supervisors felt uncomfortable about the assessment role"* and the *"reluctance supervisors have to provide critical appraisal"* (p.54). In future these issues will need addressing to outline role clarity in relation to purpose.

An additional unanswered practice question highlighted in the research was the ideal time period to be spent on supervisory activities (DECP, 1997; Hamilton-Farrell, 1993; Lunt and Sayeed, 1993; Thompson, 1993). As Lunt

(1993b) pointed out there was *"wide variation in supervision arrangements in terms of time allotted"* (p.10). Likewise 52% of newly qualified EPs felt time for supervision was insufficient (Lunt & Sayeed, 1995). The amount of time necessary for supervision to be effective needs to be explored. As Hamilton-Farrell (1993) argues *"accurate information is needed about time spent in properly defined modes of supervision"* (p.81).

All studies highlighted unresolved issues in the realm of EPiT supervision pertaining to the models of support and training for supervisors. In terms of the models of supervision, as the DECP (1987) study and others indicated, EPiTs received wide variation in the type used with a *"lack of suitable models"* (Pomerantz, 1990) (p.55). A clear practice issue is the need for a theoretical model of supervision that combines the challenging yet supportive function of educative supervision. At present the *"ambiguity regarding the purpose of supervision"* has meant no fixed or consistent model has prevailed (Nolan, 1999) (p.104). A range of the research demonstrated the need to examine a variety of models of practice from clinical models (DECP, 1987), to peer supervision (Hamilton-Farrell, 1993; Nolan, 1999) to Coaching (Lucas, 1989; Webster et al., 2000). In a similar manner all research indicated the *"relative absence of substantial training for supervisors"* (Pomerantz, 1990) (p.56). As Carrington (2004) pinpoints *"good quality training is not universally available to all supervisors"* (p.40). The issues of both models and training for effective professional supervision need further examination (DECP, 1987; Carrington, 2004; Hamilton-Farrell, 1993; Kuk & Leyden, 1993; Webster et al., 2000).

Research in the area of EPiT supervision has principally used qualitative techniques from questionnaires (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Thompson, 1999; Webster et al., 2000) to individual accounts (Carrington, 2004; Ferguson, 1990; Nash, 1999). As Pomerantz and Lunt (1993) state *"Indeed, there has been very little data to inform the profession about just how effectively it supports its members"* (p.12). Future research will need to examine this further using both quantitative and qualitative methodologies. Similarly no research has provided information on *"the reciprocal nature of the learning process in supervision"* (Carrington, 2004) (p.38). This could be relevant in ascertaining the purpose and models of supervision to be used. Finally, the benefits of supervision for the supervisor have not been adequately examined (Carrington, 2004; Hamilton-Farrell, 1993; Nash, 1999).

Section 4: Integration of Theory, Research and Practice

This section delineates the manner in which current research and practice both overlap and conflict in Educational Psychology with regard to EPiT supervision. Initially agreements and shortfalls will be explored with practice issues identified. The implications for action on both a short, medium and long term will then be outlined.

Research and Practice Issues

Some overlap exists between research and practice along key dimensions in terms of the need for EPiT supervision and the recognition of ethical practice issues. Firstly both research and practice identify the need for EPiT supervision in the profession. As Powell et al (1990) argue *"quality Supervision is not only in the interests of the trainee but also the psychological service and the LEA, not forgetting children and their parents"* (p.45). Such findings are paralleled in both the evidence from the research and consultation in the local context, with the value of EPiT supervision paramount. Indeed as Webster et al (2000) identify it is *"of crucial importance to the next generation of practitioners and to the professional body as a whole"* (p.431). Although, agreement exists that supervision is valuable, both research and practice illustrate it is an under researched entity (DECP, 1987; Hamilton-Farrell, 1993; Leyden & Kuk, 1993; Nash, 1999; Pomerantz & Lunt, 1993). Moreover, the variation in access, type and time spent on EPiT supervision is evident in practice and research domains (DECP, 1987; Hamilton-Farrell, 1993; Nolan, 1999; Thompson 1993; Webster et al., 2000). Likewise, the disparity between Educational Psychology and other disciplines is realised, which is of increasing concern given the move to integrated children's services as outlined in Every Child Matters (2003) (DECP, 1987; Lunt & Sayeed, 1995; Osborne et al., 1990; Scaife, 1993). Agreement also exists between research and practice as to the ethical considerations and potential power imbalance in supervisory relationships from LEA and BPS Guidelines (2000) to the research (Hamilton-Farrell, 1993; Nash, 1999; Nolan, 1999; Osborne, 1993).

Although there is agreement as to the relevance of EPiT supervision, research and practice conflict as to the type of supervision required. It appears a gulf exists between current practice and what research advocates as to effective supervision. In particular research evidences great disparities in the nature of supervision across LEAs with a difference between what it espouses as good practice and what is actually occurring (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Nash, 1999; Webster et al., 2000). Research details components of supervision to work effectively from the aspects of time to delineating the clarity of sessions (Carrington, 2004; DECP, 1987; Lunt, 1993b; Pomerantz, 1993; Scaife, 1993; Thompson, 1993). This espoused theory is not, however, necessarily the theory in use (Argyris & Schon, 1978). Overall research suggests consideration should be levelled at EPiT supervision, however, this is not mirrored at a practice level. Such an imbalance needs re-addressing so that the contribution of EPiT supervision can be realised at a local and national level for the continuing development of the profession.

Short Term Action to Address Practice Questions

A review of the literature identified a number of unanswered practice questions, which could be examined to ascertain effective frameworks and time required for EPiT supervision in practice. Firstly, research illustrated the controversy surrounding the dual role of EPiT supervision in both supportive and assessment functions. As Lunt, (1993b) indicates there is *"a lack of clarity over the purposes of supervision, or indeed disagreement over its purpose"* (p.11). In particular the discomfort many supervisors experienced

in critically appraising whilst at the same time supporting EPiTs was indicated (Carrington, 2004; Pomerantz, 1990; Powell et al., 1990). At an LEA level this combined role needs careful consideration and clarity so that both supervisor and supervisee have a joint understanding of both roles and functions. Such work could be kick-started by an audit and consultation within and between LEA's and HEI's to develop clear definitions of EPiT supervision modes and types drawing on BPS/ DECP Quality Standards (2000; 2002). This liaison between HEI's and LEA's could be crucial particularly given the imminence of extended training (Malcolm, 2003; Turner, 2003). Consultation and the use of theoretical models and research could then translate to policy to affect practice. Moreover separating out *"the learning process from the assessment process"* could be detailed (Thompson, 1993) (p.89).

At a policy level the prominence of supervision could be increased at an LEA level in terms of managing change (Leyden & Kuk, 1993). Creation of specialist posts linked to supervision particularly Practice Tutors and Field Supervisors would need specifying (AEP, 2004a; Malcolm, 2003; Turner, 2003). As the AEP (2004b) notes *"such supervisors will have additional responsibilities and duties, which will need to be recognised both by salary and status"* (p.2). The creation of such posts will mark the increasing importance of EPiT supervision over an extended training period within LEA contexts. Alongside this, training for supervisors will be required within LEA's and HEI's (Carrington, 2004; DECP, 1987; Powell et al., 1990; Turner, 2003). Such investment will mean the competencies a practitioner requires in order

to provide high quality supervision are addressed (Falender et al., 2004). Supervisory practices will, likewise, need review in the wider context given the onset of integrated children's services within the LEA setting as espoused in Every Child Matters (H.M.Treasury, 2003). In this manner Educational Psychology will need to ensure its supervisory structures are comparable to fellow professionals, which at present is not the case (Lunt, 1993b; Lunt & Sayeed, 1995; Pomerantz, 1993b). As a discipline, supervision will need to be embedded in service protocols in order that EPs are recognised as rigorous and reflective practitioners alongside colleagues in multidisciplinary teams.

An additional unanswered practice question was the amount of time required for effective supervision as this varied considerably across LEAs (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Thompson, 1993; Webster et al., 2000). Accurate information regarding time for phases of supervision is needed (Hamilton-Farrell, 1993) with time protected in the LEA context (Pomerantz, 1993a). Such matters will need recognition between LEA's and HEI's to stipulate best practice at different phases of EPiT training (Malcolm, 2003; Turner, 2003). Such discussion should inform policy and practice to ensure time is protected (AEP, 2004a). This should also be the case in terms of established practitioners with supervision established as part of the continuing professional development programme with time allocated accordingly (Lunt, 1993b).

Medium Term Action to Examine Unresolved Issues

A number of unresolved issues were left unexamined in the literature, which could be explored through research and practice. A key area was the lack of clarity surrounding supervisory models. This could be addressed through LEA's and HEI's reviewing practice and research to generate curriculum changes in the supervisory frameworks as necessitated by extended training (Malcolm, 2003; Turner, 2003). As Webster et al (2000) highlight this could involve the development of a professional curriculum for supervision which focuses on *"systematic, reflective, evidenced-based practice"* (p.433). LEA's and HEI's could co-operatively work together in the construction and support of a curriculum with *"new conceptual and functional frameworks for EPiT supervision"* (Webster et al., 2000) (p.432). Indeed the need to explore alternative supervisory frameworks has been illustrated in a range of work (DECP, 1987; Lucas, 1989; Osborne et al., 1990; Powell et al., 1990).

LEA's and HEI's could examine different models of supervision to ascertain a best fit for EPiT support (Osborne et al., 1990; Powell et al., 1990). Given the changes in training a developmental approach could be explored for as Powell et al., (1990) indicate, *"a supervisee's needs change over time"* (p.49). This would be particularly pertinent over an extended training period and research advocates a need for such a developmental approach (Carrington, 2004; Hamilton-Farrell, 1993; Powell et al., 1990). Such a model would ensure supervision was responsive and flexible to individual EPiTs level of progress over time (Nash, 1999). One such model to be investigated could be the Counsellor Complexity Model (Stolenberg, 1981), which

identifies four levels for supervisee and supervisor in terms of learning. As Stolenberg (1981) outlines *"the delineation of these levels of development and the concomitant environments for facilitatory development through the levels will provide a useful model for conceptualising the supervision process"* (p.64). Indeed such an approach would ideally delineate phases of supervision throughout an EPiT's evolution alongside that of established practitioners.

An additional approach to be examined could be the *"seven eyed supervisor model"* (Proctor & Inskipp, 1991). This differs from other models of supervision as it looks specifically at the process of the relationship alongside the context and wider organisational issues. Such psychotherapeutic and systemic models could be examined to extend the scope of supervision (Osborne et al., 1990; Powell et al., 1990). These clinical models could be complimented by alternatives such as Consultation (Caplan, 1970) and Coaching (Kearney, 1995) to allow for equal participation of supervisee and supervisor (Lucas, 1989; Osborne et al., 1990; Webster et al., 2000). In particular these are regarded as critical in developing reflection in *"flexibly supportive frameworks for mutual self-evolving"* (Lucas, 1989) (p.47). Finally peer group supervision could be examined as *"group supervision can offer a rich tapestry for learning and development"* (Scaife, 1993) (p.64). Exploration could occur into the different types of supervision group from co-operative supervision to peer group supervision as defined by Proctor and Inskipp (1991). This is an area worthy of future investigation given the mutual support and development it provides (Hamilton-Farrell, 1993; Zorga et al.,

2001). All such models could be examined in research and practice between LEA's and HEI's alike to construct consistent frameworks.

Another unresolved issue identified in research and practice was the need for supervisor training. As Carrington (2004) suggests there is a *"need for more attention to be given to the training of supervisors"* (p.40). Indeed Nolan's 1999 study indicated that supervisors *"thought that specific training would be beneficial"* (p.103). As Webster et al (2000) indicate the *"interface which training institutions must work towards with LEA's is in order to implement practice orientated training" (*p.432). The need for such training in EPiT supervision and for consistency has been raised (Carrington, 2004; DECP, 1987; Osborne et al., 1990). One means of achieving this could be the development of a core curriculum for supervision training across LEA's and HEI's building on the original work of the BPS Guidelines on Fieldwork Supervision (1990). As Powell et al (1990) argue *"a common curriculum for supervision training would contribute to a greater consistency of supervision received by trainees"* (p.45). This could entail work preparing both EPiTs, Field Tutors and Supervisors for the processes and varying functions of supervision. As Powell et al (1990) state *"course tutors help prepare the supervisors through the medium of supervision training and also prepare the trainees to be receptive to and active seekers of the fruits of such training"* (p.50). Moreover, in the LEA context the Practice or Field Tutor Co-ordinator role could establish collective supervision for practitioners supervising EPiTs. It is argued that *"services and courses should consider the question of how*

best to create supervisory opportunities for the supervisors" (Powell et al., 1990) (p.50).

Given the advent of extended training within the profession, it is apparent that the training of supervisors will need to be addressed from a combined approach from the LEA's and HEI's (Malcolm, 2003; Turner, 2003). As Kuk and Leyden (1993) point out *"services and training courses could combine to provide in-house training and development work in supervision for the benefits of both parties"* (p.60). The opportunities for this could occur with individual LEA's or services nationally to create consistent training for supervision. As Webster et al (2000) argue *"these links – between EPiTs, LEA services and training courses – are a source of collective professional strength and might well become an even more important triangle in the future"* (p.447).

Long Term Action to Explore Future Research Directions

A review of the psychological literature highlighted a range of research possibilities within the domain of EPiT supervision, which were omitted but would be informative to practice. The majority of studies were limited by inadequate qualitative designs with a restricted focus. As Hamilton-Farrell (1993) argues *"there is clearly a need for further investigation of fieldwork Supervision"* (p.80). Indeed, to date research has primarily relied on questionnaire surveys (DECP, 1987; Hamilton-Farrell, 1993; Lunt & Sayeed, 1995; Thompson, 1999; Webster et al., 2000) or individual practitioner accounts (Carrington, 2004; Ferguson, 1990; Nash, 1999). Pomerantz,

(1990) details how *"we obviously require more recent and up-to-date research published in our field"* (p.54). One way forward would be to improve and enhance the research design used, particularly the qualitative methodologies utilised with a clear rationale specified. In terms of measurement devices the development of reliable and valid tools to tap into perceptions could be created with the relevant technical details and standardisation. Moreover, more research could combine both quantitative and qualitative methodologies. The use of larger samples and specified demographic characteristics could support the control of confounding variables. In this fashion, research could be significantly improved by scientifically rigorous designs to allow for credible and replicable outcome results. This would address the paradox in the profession given the consensus as to the importance of supervision yet variable information as to its implementation and impact.

Furthermore, the scope of EPiT supervision could be extended within the realms of research with LEAs and HEIs alike. Firstly the supervisory relationship could be focused in on to illuminate the variety of mechanisms at work. As Stolenberg (1981) indicates *"the factors appearing to be most instrumental in effecting change during the supervision process can be operationalised and subjected to empirical scrutiny"* (p.64). The actual learning process inherent in such a relationship could, likewise, be explored (Carrington, 2004; Thompson, 1999). Similarly, the potential benefits of such learning could be ascertained for as Carrington (2004) pinpoints, *"the potential learning gains of supervising remain largely unexplored and unlike*

the benefits of receiving supervision, generally un-researched" (p.34). The actual phenomena of supervision could also be gauged as *"more information is needed about the practice and experience of both trainees and supervisors"* (Hamilton-Farrell, 1993) (p.80).

Research into EPiT supervision could examine alternative and varying models to reveal potential effectiveness. In particular a need exists to test models empirically in supervision with people in a given profession (Stolenberg, 1981). This is pertinent given extended training will necessitate different supervisory frameworks to accommodate for the developing skills and experience of EPiTs over time. As Pomerantz (1990) states, *"it would help if fieldwork supervision became the subject of rigorous discussion and debate in order to highlight the rich diversity of methods and ideologies which are available"* (p.54). Supervision models could be compared in an alternative intervention design in LEA and HEI contexts to consider applicability through the extended training. For example, group supervision could be compared to individual supervision given it economies in time and money and enhanced benefits of extended experience (Hawkins & Shohet, 2002). Similarly, medical models could be contrasted to those of Consultation and Coaching (DECP, 1987; Lucas, 1989; Thompson, 1993; Webster et al., 2000). As Nolan (1999) highlights *"the development of models for EPs could be developed with practising EPs and researched over time"* (p.106). Such work would help to delineate the different types of supervision required throughout EPiTs training. This research would need to occur within the profession as *"the application of expertise is linked to*

enculturation within a profession. This fact is often neglected in expert research" (Bromme & Tillema, 1995) (p.264). This could examine the extent to which supervision is embedded within the discipline. Future avenues of enquiry need to be actively pursued for as Osborne et al (1990) indicate *"despite the importance of fieldwork supervision in the training of Educational Psychologists, there has been little previous study of this aspect"* (p.38). Such an issue needs resolving with joint HEI and LEA research into effective and efficient supervision.

Concluding Comments

This paper demonstrated the increasing importance of EPiT supervision at a local and national level. The literature review parallels such interest with a range of research examining present practice. Despite agreement as to the value of EPiT supervision, substantial variation exists, however, in context. The need to address shortfalls between research and practice are illustrated by short, medium and long term action. The implementation of consistent and comparable EPiT supervision across domains is imperative to ensure quality practice for future generations. This is particularly pertinent given the onset of extended training for as Webster et al., (2002) highlight *"this is an opportune time for the profession to rethink its nurturing role"* (p.447).

References

Argyris, C. and Schon, D. (1978). *Organisational Learning*. Reading Mass: Addison Wesley.

Arnold, J., and Johnson, K. (1997). Mentoring in Early Career. *Human Resource Management*, 4, 61-70.

Association of Educational Psychologists. (2004a). *Initial 3-Year Doctorate Training for Educational Psychologists, An Update.* Circular 026/04, 2nd April 2004.

Association of Educational Psychologists. (2004b). *A Salary and Service Condition Claim*, 20th June, 2004.

Belsky, J. and Fearon, R.M.P. (2002). Early Attachment Security, subsequent maternal sensitivity, and later child development: Does continuity in development depend upon continuity of care giving? *Attachment and Human Development*, 4(3), 361-387.

Binsted, D. (1980). Design for learning in management training and development: a view. *Journal of European Industrial Training*, 4 (8), 1-32.

British Psychological Society. (1987). *Fieldwork Supervision.* Report of the Joint DECP Training Committee/Course Tutors Working Party, Leicester: The British Psychological Society.

British Psychological Society. (1990). *Guidelines on fieldwork supervision, Leicester*: BPS.

British Psychological Society Division of Educational and Child Psychology. (2000). *Quality Standards for Educational Psychology Services.* Leicester: BPS.

British Psychological Society Division of Educational and Child Psychology. (2002). *Professional Practice Guidelines.* Leicester: BPS.

Bromme, R. and Tillema, H. (1995). Fusing experience and theory: the structure of professional knowledge. *Learning and Instruction*, 5, 261-267.

Buchanan, D. and Huczynski, J. (1997). *Organisational Behaviour.* (3rd edition). Prentice Hall, Hemel Hempstead.

Caplan, G. (1970). The theory and practice of mental health consultation. *Consultee-centred consultation*. New York: Basic Books.

Carrington, G. (2004). Supervision as a Reciprocal Learning Process. *Educational Psychology in Practice,* 20 (1), 31- 42.

Department for Education and Employment. (2000). *Educational Psychology Services (England): Current Role, Good Practice and future Directions.* Report of the Working Group, London: DfEE.

Division of Educational and Child Psychology. (1987). *Fieldwork Supervision.* Report of the Joint DECP Training Committee/Course Tutors Working Party, Leicester: The British Psychological Society.

Dowling, E. and Osborne, E. (1994). *The Family and the School – A Joint Systems Approach to problems with Children.* (2nd Edition). London: Routledge.

Falender, C.A., Erickson Cornish, J.A., Goodyear, R., Hatcher, R., Kaslow, N.J., Leventhal, G., Shafrankse, E., Sigmon, S.T., Stoltenberg, C., and Grus, C. (2004). Defining Competencies in Psychology Supervision: A consensus Statement. *Journal of Clinical Psychology*, 60(7), 771-785.

Fenwick, T.J., (2003). Professional growth plans: Possibilities and Limitations of an Organisation wide Employee Development Strategy. *Human Resource Development Quarterly*, 14(1), 59-77.

Ferguson, D. (1990). So you're going to supervise an Educational Psychologist in Training? *Educational Psychology in Practice*, 6(3), 131-135.

Gersch, I.S. (2001). *Educational Psychology: preparing for the future.* Presentation at Surrey Educational Psychology Service Conference, 27th March 2001.

Hamilton-Farrell, J. (1993). Experience of supervision from the perspective of educational psychologists in training. *Educational & Child Psychology*, 10(2), 73-81.

Hawkins, P. and Shohet, R. (2002). *Supervision in the helping professions*. Milton Keynes: Open University Press.

H.M. Treasury. (2003). *Every Child Matters*. (Cm 5860). TSO.

Jennings, C. (1995). Challenges and Opportunities: future training perspectives and the changing roles, tasks and contexts in educational psychologists' work. *Educational and Child Psychology*, 12(2), 15-24.

Joiner, T.A., Bartram, T., and Gareffa, T. (2004). The Effects of Supervision on Perceived Career Success, Commitment and Turnover Intentions. *The Journal of American Academy of Business*, September, 164-170.

Kearney, D. (1995). *Coaching*. Yeovil, Somerset: Organisational and Personal Development Consultants.

Kleinman, G., Siegel, P.H., and Eckstein, C. (2001). Mentoring and learning. *Leadership and Organization Development Journal*, 22(1), 22-34.

Kuk, G., and Leyden, G. (1993). "What's in it for us?" Supervision and educational psychologists: analysis of survey returns. *Educational and Child Psychology*, 10(2), 51-60.

Leyden, G., and Kuk, G. (1993). The role of Supervision in a Healthy Organisation: The Case of Educational Psychology Services. *Educational and Child Psychology*, 10(2), 43-50.

Lucas, P. (1989). Clinical Methods of Supervision: a note of caution and challenge. *New Era In Education*, 70(2), 42-47.

Lunt, I. (1993a). Editorial. *Educational and Child Psychology*, 10(2), 3.

Lunt, I. (1993b). Fieldwork supervision for Educational Psychologists in training. The 1984-5 survey and the State of the Art at That Time. *Educational and Child Psychology*, 10(2), 4-11.

Lunt, I., and Sayeed, Z. (1995). Support for educational psychologists in their first year of professional practice: induction and supervision arrangement for newly qualified educational psychologists. *Educational and Child Psychology*, 12(2), 25-30.

Malcolm, L. (2003). *New Entry Training Route for Educational Psychologists (England): Report of an Implementation Study carried out on behalf of the Educational Psychology Training and Development Subgroup.* Applied Psychologists Associates Ltd.

Nash, J. (1999). Developing and Refining Supervisory skills. *Educational Psychology in Practice,* 15(2), 108-115.

Nolan, A. (1999). Supervision for Educational Psychologists: how are we doing? *Educational Psychology in Practice*, 15(2), 98-107.

Orpen, C. (1997). The effects of formal mentoring on employee work motivation, organizational commitment and job performance. *The Learning Organisation*, 4(2), 53-60.

Osborne, E., Leyden G., and Powell, M. (1990). Supervision of trainee educational psychologists on fieldwork. *Educational and Child Psychology*, 7(3), 37-43.

Osborne, E. (1993). External Pressures and Influences on Face-To-Face Supervision. *Educational and Child Psychology*, 10(2), 25-30.

Pomerantz, M. (1990). Fieldwork Supervision: A Personal Perspective On the State of the Art. *Educational and Child Psychology*, 7(3), 53-58.

Pomerantz, M. (1993a). The practice of supervision for local authority educational psychologists. *Educational and Child Psychology*, 10(2), 16-24.

Pomerantz, M. (1993b). The value and purpose of supervision for educational psychologists. *Educational and Child Psychology*, 10(2), 31-34.

Pomerantz, M., Leyden G., Lunt, I., Osborne, E., Powell, M., and Ronaldson, J. (1987). *Fieldwork supervision: Report on the joint DECP Training Committee/Course Tutors' Working Party*, Leicester: British Psychological Society.

Pomerantz, M. and Lunt, I. (1993). Investigating Supervision Practice: History and Methodology of the Enquiry. *Educational and Child Psychology*, 10 (2), 12-15.

Potter J. and Wetherell, M. (1987). *Discourse and Social Psychology. Beyond Attitudes and Behaviour*. Sage Publications.

Powell, M., Leyden G., and Osborne, E. (1990). A curriculum for training in supervision. *Educational and Child Psychology*, 7(3), 44-51.

Proctor, B. and Inskipp, F. (1991). Group Supervision. Chapter 6 in P. Hawkins and R. Shohet, (2002) *Supervision in the helping professions*. Buckingham: Open University Press.

Ramchandi, P., Joughlin, C. and Zwi, M. (2001). Evidence based child and adolescent mental health services: Oxymoron or brave new dawn. *Child Psychology and Psychiatry Review*, 6, 59-64.

Rudestam, K.E. and Newton, R.R. (1982). *Surviving your Dissertation.* Newbury Park.

Sayeed, Z.N. and Lunt, I. (1992). Induction and supervision for newly qualified educational psychologists, *Educational Psychology in Practice*, 8(3), 156-164.

Scaife, J. (1993). Application of a general supervision framework: creating a context for supervision. *Educational and Child Psychology*, 10(2), 61-72.

Scaife, J. (2001). Supervision in the Mental Health Professions. *A Practitioners Guide.* Bruner Routledge.

Schon, D. (1991). *The reflective practitioner*: *How professionals think in action.* Aldershot: Avebury, Ashgate Publishing Ltd.

Sherrard, C. (1997). Qualitative Research. *The Psychologist*, 10, 161-162.

Stolenberg, C.D. (1981). Approaching supervision from the developmental perspective. *Journal of Counselling Psychology,* 28, 61-66.

Surrey Children's Service. (2004). *Supervision Development*. Surrey County Council.

The Tavistock and Portman NHS Trust. (2003). *Supervision: A Psychological Approach.* Course D57.

Thompson, D. (1993). Dilemmas In Evaluating Competence On Practical Placement. *Educational and Child Psychology*, 10(2), 82-29.

Turner, J. (2003). *Restructured Professional Training for Educational Psychology (England). Update for NAPEP Conference*. Presentation at the NAPEP Conference, 30th September 2003.

Vecchio, R.P. and Boatwright, K.J. (2002). Preferences for idealized styles of supervision. *The Leadership Quarterly*. 13, 327-342.

Webster, A., Hingley P., and Froney, J. (2000). Professionalization and the reduction of uncertainty: a study of new entrants to educational psychology. *Educational Psychology in Practice*, 16(4), 431-448.

Willig, C. (2001). *Introducing Qualitative Research in Psychology. Adventures in Theory and Method.* Open University Press.

Zorga, S., Dekleva, B and Kobolt, A. (2001). The process of internal evaluation as a tool for improving peer supervision. *International Journal for the Advancement of Counselling,* 23, 151-162.

Professional Practice Assignment 4: Problem Based Learning: The Panacea for Professional Training and Transformation of Tutor Role

Abstract

Society now demands professionals who are not only competent but have the capability to develop and translate the existing knowledge base to varying contexts in a team setting (Dochy et al., 2003; Fraser & Greenhalgh, 2001; Henlow & Evensen, 2000; Macdonald & Savin–Baden, 2004). Consequently the professional training of educational psychologists has a responsibility to meet such requirements. Problem Based Learning (PBL) offers a vehicle for achieving all such demands whilst managing curriculum overload from a discipline in its infancy. This paper examines the PBL literature with particular regard to the role of tutor. Research highlights numerous issues and questions for consideration which are used to delineate the best direction forward. As Schwartz (2004) highlights "*it is time to move away from passive to active forms of learning- from absorbing facts to solving problems*" (p.26).

Section 1: Aims and Scope of Assignment

Interest in PBL as a technique has been apparent in Higher Education Institions (HEIs) and is acknowledged as the method of choice in professional training from medicine to law (Macdonald & Savin-Baden, 2004; Newman, 2004; Schwartz et al., 2001). Furthermore, the imminence of extended training in educational psychology has lead to its inclusion on courses in both Dundee and Manchester (EPiT Conference, 2005; Kerr &

Smith, 2005). It appears PBL will act for many universities as the means of redesigning and delivering future curriculum conceptualisations.

This paper aims to explore one crucial aspect of PBL, namely the transformation of tutor. Such a specific focus was adopted because as Hmelo-Silver (2004) argues "*the facilitator role is critical to making PBL function well*" (p.244). The importance of the tutor is evident (Barrows, 1986; Hmelo-Silver, 2004; Maudsley, 2002) as are the requisite changes in role requirements, from tutor as disseminator of content to facilitator (Elton, 2004; Hmelo-Silver, 2004). The potential challenges this generates have been documented and as such represent an area worthy of exploration (Boud & Feletti, 1997; De Weerdt et al., 2002; Houlden et al., 2001; Macdonald & Savin-Baden, 2004; Schwartz et al., 2001; Tohey, 1999; Vernon, 1995). As present "*the literature provides little insight into PBL tutors' collegial (dis) harmony about their role*" (Maudsley, 2002) (p.163) and acts as an area for further investigation for this paper. A more extensive study of PBL did not occur given the scope of this piece.

Definitions

> *"Problem Based Learning (PBL) is an approach to professional education that makes use of real-life problems as a stimulus for learning. In PBL students work in small tutorial groups on these problems and in the course of discussing them, formulate goals for self-directed learning."*
>
> (Van Berkel & Schmidt, 2000) (p.231).

The technique begins with a problem which students then focus on in terms of their existing pooled knowledge and areas for investigation. The tutor acts as a source of facilitation in terms of both the content and the process of learning (Macdonald & Savin-Baden, 2004). The group works through hypotheses and then return later to apply and reflect upon them in the given scenario. The approach can vary for as Barrows (1986) highlights "*the term problem based learning must be considered a genus for which there are many species and subspecies*" (p.485). As Walton and Matthews (1989) state, 3 core features permeate all PBL programs. These include **curricular organisation** around problems not disciplines; the **learning environment** as a small group tutorial with self-directed learning; and **stated learning outcomes** which encourage the development of skills and life long learning. Such core factors are espoused and detailed in full in Boud, 1985; Dochy et al., 2003 and Newman, 2004.

Overview

This assignment presents an overview of the rationale and relevance of PBL in the local and national context. A critique of the research will follow with a focus on the changing role of tutor given its importance in the success of the strategy (Barrows, 1986; Hmelo-Silver, 2004; Maudsley, 1999). Such evaluations will then inform practice applications and outline future research directions in professional training.

Section 2: Practice and Context

The rationale for examining PBL as a future tool for professional training in educational psychology exists at both a local and national level.

Local Context: Development of PBL in the University Setting

Establishments are facing changing demands given the extended training of educational psychologists which warrant consideration (BPS, 2004). Moreover increasing pressure from society pushes for teaching not only content but, to develop skills in applying and modifying this knowledge to changing scenarios (Dochy et al., 2003; Fraser & Greenhalgh, 2001; Henlow & Evensen, 2000; Stephenson & Yorke, 1995). A corresponding shift in training is needed as Schwartz (2004) highlights "it *is time to change the way we teach psychology*" (p.26). This is particularly pertinent given the expanding psychological knowledge base. Such factors mean students need support in how to manage and update new knowledge as opposed to merely possessing it (Dochy et al., 2003; Henlow & Evensen, 2000; Savin-Baden, 2004). Professional training has to, therefore, meet the dual demands of transmitting content whilst evolving practice skills of application in the context of ever growing theory. PBL promises a powerful device to achieve such objectives, being responsive to developments in research and yet compatible with societal need (Dochy et al., 2003; Newman, 2004). This may explain why "*PBL is becoming widely used across a considerable range of subjects and professional areas in higher education*" (Macdonald & Savin-Baden, 2004) (p.2).

Educational Psychology training course staff are becoming increasingly interested in the use of PBL as a tool for curriculum evolution (EPiT Conference, 2005; Kerr & Smith, 2005). Central to this is the changing face of the tutor in shaping and implementing such developments. As Barrows (1986) indicates "*a potent effect on the quality of PBL, is the understanding and skills of the teacher or tutor*" (p.245). This has been indicated by work in the field (Dundee EPiTs, 2005; Hmelo-Silver, 2004; Kerr & Smith, 2005; Maudsley, 2002). In PBL the tutor role changes from teacher to student centered, and from teaching to facilitation (Elton, 2004). Hmelo-Silver (2004) depicts how "*the teacher acts as a facilitator to guide student learning through the learning cycle*" (p.236). Such a shift in both skills and attitudes is an issue which universities will need to manage. As Macdonald and Savin-Baden (2004) describe "*it is very difficult for some tutors to let go and to stop providing answers to the questions when those very questions provide the basis for students' learning activities*" (p.4). Indeed adjustment difficulties from tutor groups have been evident (Houlden et al., 2001; Ryan, 1997; Toohey, 1999; Vernon, 1995). In particular the tension of the tutors' position in providing the right level of intervention whilst acting as a role model of reasoning has been highlighted (De Weerdt et al., 2002). Such issues have been reported as a challenge for tutors (Boud & Feletti, 1997; Ryan, 1997; Savin-Baden, 2000) with subsequent resistance (Mennin & Kaufman, 1989; Schwartz et al., 1994; 2001; Thompson & Williams, 1985). Given the importance of the tutor in the effectiveness of PBL this is an area in need of exploration (Boud & Feletti, 1997; Maudsley, 2002; Savin-Baden, 2000; Schwartz et al., 2001).

Finally at the local level of universities PBL presents a further method to support the application of adult learning models intrinsic to professional training (Cline et al., 1990; Frederickson et al., 2003; Monsen, 2000; Wright, 1990). The technique takes its foundations from a range of cognitive models linked to teaching and learning (Barrows, 1994; Newman, 2004; Norman & Schmidt, 2000). These range from the social construction of knowledge (Vygotsky, 1978) to the need for reflection (Schon, 1987) and self-directed learning (Hmelo & Lynn, 2000). Furthermore they build on adult learning models (Kolb, 1984) and the process of evolving theory in use (Argyris & Schon, 1974). In this manner PBL is in keeping with existing strands of practice in universities in the training of practitioners which, are derived from the psychological theory base.

National Context: Development of Training to Meet Societal Demand

On a national level the rationale for this assignment originates from two forces, principally the evolution of training to meet workforce requirements and to assist in the progression of the profession.

> *"In today's complex world, we must educate not merely for competence, but for capability (the ability to adapt to change, generate new knowledge, and continuously improve performance)"*
>
> (Fraser & Greenhalgh, 2001) (p. 799).

Such is the challenge for educators today from the employment market. However, as Dochy et al., (2003) argue "*educational practices have been criticized for not developing these prerequisites of professional education*"

(p.534). The issue is now how to ensure an education which develops students' skills in applying and developing knowledge effectively. Alongside this practitioners are required to deploy such skills in a team setting (Henlow & Evensen, 2000). Recent legislation echoes such principles with multi-professional collaboration at the heart of Children's Services (DfES, 2004). Subsequently professional training of educational psychologists will need to encompass such goals. In particular the emphasis on multi-agency working with fellow professionals will need addressing (NAEIAC/AEP, 2005). This is recognised; "*investment in team learning would seem to us to be a vital component of higher education – one that we should embrace – and problem based learning can help us do just that*" (Macdonald & Savin-Baden, 2004) (p.3). PBL offers a solution to teach both content knowledge and skills based practice in a team setting (Barrows, 1994). As Schwartz et al (2001) highlight "*it is a method of learning in which learners first encounter a problem, followed by a systematic, student centered enquiry process*" (p.1/2). In this manner it encapsulates the professional artistry needed in translating theory into practice in a seamless manner (Lunt & Majors, 2000) as well as embedding it in a team context as warranted by workforce demand (Dochy et al., 2003; Fraser & Greenhalgh, 2001; Henlow & Evensen, 2000).

At a statutory level the imminence of extended training provides an additional reason for pursuing PBL in Higher Education Institutions (HEIs) (BPS, 2004). The DECPTC accreditation criteria attempts to align educational psychology training to other strands of the science. This is in accordance with European standards for professional training in applied psychology and the draft Quality

Assurance Agency (QAA) benchmark standards which require all chartered psychologists to have doctoral level award. PBL offers an approach which will assist educational psychology to extend training so it resembles other applied branches of the profession. In addition the technique meets the DECPTC accreditation criteria with its emphasis on self-directed learning skills. Indeed PBL places the learner as a central resource and feature of the process (Hmelo & Lynn, 2000).

The impetus for improving the professional training of educational psychologists exists at both a local and national level. PBL provides a vehicle to both build and extend current practice whilst simultaneously meeting the demands of the 21st century workplace. In order to successfully implement such a strategy the role of tutor merits closer examination given its centrality. This assignment seeks to explore how this could occur for the future advancement of the discipline.

Section 3: Psychological Theory and Research

The literature search used the bibliographic data bases of Medline, Psyclit, PsychInfo and Educational Resources Information Centre (ERIC) as recommended by Ramchandi et al (2001). Key search terms utilised were *"problem based learning",* and *"tutor role in PBL."* Manual searches occurred in conjunction with ancestral searches (Smits et al., 2002).

The review aims to examine the research surrounding the tutor role in PBL. The studies included used problem based interventions originating from the methods proposed by Harvard Medical School (Barrows, 1986). Research was included from the professional domains of Health and Education as these are most akin to educational psychology. The studies selected are detailed overleaf in an overview. The critique uses existing research as reported in the field (Colliver, 2000; Hmelo-Silver, 2004; Macdonald & Savin-Baden, 2004; Newman, 2004; Norman & Schmidt, 2000; Smits et al., 2000; Van Den Bosse et al., 2000). It incorporates studies involving quasi-experimental design to ensure ecological validity to the area of examination. In this manner internal validity relative to more controlled experimental environments may be forfeited.

The critique is divided into 5 sections using the critical evaluation checklist of Rudestam and Newton (1982) alongside guidance from Greenhalgh (2001) and Connor (1997). This encompasses a theoretical, empirical, practical and ideological critique (Broughton, 1981). In addition the studies are examined according to Friedman et al's 1990 essential characteristics of evaluation studies in curricula change. Initially theoretical frameworks and conceptualisations will be examined. Following on from this the empirical design is critiqued alongside the results and discussion. A final section highlights the key themes and controversies.

AUTHORS	FOCUS OF STUDY	THEORETICAL CONCEPTUALISATIONS	RESEARCH DESIGN	RESULTS AND DISCUSSION
Premi et al (1994)	Preliminary evaluation after the first year from students and staff to ascertain attitudes toward PBL.	No research questions detailed, but reference made to earlier studies. Processes of PBL detailed but no distinct model specified.	Pre and post questionnaires to evaluate PBL program over 6 months. Insufficient detail as to qualitative design and interpretation.	Positive responses to PBL from both groups but lack of detail as to analysis. Limitations of design considered but no alternative conclusions considered.
Bernstein et al (1995)	Evaluation of shifts in students' attitudes and to describe faculty experiences of PBL.	Research questions and rationale for study linked to empirical evidence. No explicit PBL model or principles referred to.	Pre and post questionnaires over a 6 week period to examine attitude shift. Procedures of qualitative method omitted alongside variables of investigation.	Direct experience of PBL led to more favourable attitudes. Qualitative interpretation of data not addressed. Alternative explanations and limitations omitted.
Kaufman and Holmes (1996)	Examination of student and staff perceptions during the first two years of transition to PBL.	Clearly stated research questions with rationale. Explicit reference to the COPS approach (Case Orientated Problem Stimulated Learning).	Post PBL questionnaire with detailed variables and defensible category choices. Design clear in terms of sample and variables used.	Tutors and students satisfied. A need for training was identified. Limitations of design not discussed nor were practical implications.

Wilkerson and Maxwell (1988)	Examination of the characteristics and beliefs of staff who volunteered as tutors for PBL teaching.	No research questions but study's purpose contextualised in earlier work. No clear model or practice for PBL specified.	Semi-structured interviews and questionnaire survey after a 2 year transition to PBL. Use of existing evidence to explain design.	Tutors interested in developing PBL tended to be interested in teaching and reform. Practical steps considered for future use.
Vernon (1995)	A description of the attitudes and opinions of tutors in PBL programs across a range of university settings.	No research questions but reference made to earlier work. Clear models of PBL and processes delineated.	Questionnaire following PBL experience from large sample of 822 tutors from 22 universities. Detailed design in terms of methods used.	Overall tutors preferred PBL to traditional approaches. Limitations and qualitative interpretation discussed. Practical issues considered.
Houlden et al (2001)	An analysis of PBL tutors' experiences to identify their problems in facilitating the curriculum.	No research questions or rationale for study provided. Reference made to a hybrid model of PBL yet no detail.	Semi-structured interviews used to identify tutor issues. Limited rationale for qualitative design adopted.	Key problems linked to students in PBL. Qualitative data discussed as were next steps. Limitations omitted.
Maudsley (2002)	An exploration of a first cohort of tutors in making sense of PBL.	No research questions yet rationale for research connected to earlier work. Detailed description of theory and processes of PBL.	Semi-structured telephone interviews with tutors after 1 years experience of PBL. Justified use of design detailed.	Positive response to PBL reported. Qualitative analysis and interpretation detailed. Limitations and implications tackled in depth.

1. Conceptualisation

Initial studies in the PBL literature essentially examined the effectiveness of the approach in relation to students alone. The exclusive focus of research concentrated on changes in learners' opinions and knowledge (Albanese & Mitchell, 1993; Berkson, 1993; Martenson et al., 1992; Vernon & Blake, 1993). An overview of research in the area called for "*more studies of faculty and practitioner satisfaction since most schools are already addressing student satisfaction*" (Friedman et al., 1990) (p.13). Subsequently work started to evaluate both student and staff perspectives. Research from Premi et al., (1994), Bernstein et al., (1995) and Kaufman and Holmes (1996) explored shifts in attitudes of both groups following direct experience of PBL. Work, likewise, became focused on the role of tutor in particular as "*little systematic attention to the opinions of faculty in evaluating problem based learning* "was given (Vernon, 1995) (p.216). In this manner Vernon's 1995 study focused on tutor's attitudes toward PBL in comparison to traditional methods and investigated influences on these perceptions. Similarly Houlden et al (2001) explored the difficulties new tutors encountered in facilitating a PBL curriculum whilst Wilkerson and Maxwell (1988) examined tutors who volunteered for PBL teaching. Finally Maudsley (2002) examined how a cohort of tutors characterised the new PBL curriculum and made sense of the learning involved. Consistency among the studies is evident in that attitudes to PBL are the main variable of interest as a teaching/tutoring technique in professional education.

Overall the processes and procedures of PBL were delineated in the context of the university under study (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988). A number of studies, however, did not explicitly explain the specific PBL model utilised. Indeed only Vernon (1995) and Kaufman and Holmes (1996) make direct reference to a certain PBL model. In terms of theoretical underpinnings this indicates a considerable omission which impacts on credibility. In fact only one piece made clear connections to the philosophy underlying PBL and its links to cognitive psychology (Maudsley, 2002). This resonates with Newman's (2004) comments that as there was a "*lack of detailed information in the papers included it is difficult to distinguish where PBL is used*" (p.22).

The majority of research exploring the tutor role in PBL has been original and has built upon itself over time (Greenhalgh, 2001). Premi et al (1994) were among the first to evaluate student and staff perceptions as opposed to educational outcomes. Bernstein et al (1995) developed this by using a more scientifically rigorous analysis of qualitative and quantitative data. The length of studies were then extended with Kaufman and Holmes (1996) examining a two year transition period to PBL. Research also honed in on tutors exclusively with Wilkerson and Maxwell (1988) concentrating on this group. This was justified in relation to the empirical evidence base with "*virtually no research on motives, characteristics, and perceptions of individual adopters*" (Wilkerson & Maxwell, 1988) (p.892.) Similarly, research was expanded by Vernon (1995) who examined tutors perspectives from a number of

universities rather than a single department. Later studies utilised more rigorous research designs in connection to tutor's opinions. Houlden et al (2001), for example, used grounded theory (Strauss & Corbin, 1990) whilst Maudsley (2002) used semi-structured interviews to compliment questionnaire surveys. In summary much cumulative development is apparent in this area with later studies often building on earlier ones to offer insight into the role of tutor in PBL.

2. Theoretical Framework and Hypotheses

With regards to theoretical frameworks the research varied. A few studies effectively and explicitly referred to the principles of PBL and the models pertaining to this (Kaufman & Holmes, 1996; Vernon, 1995). One study in particular detailed the psychology of adult learning with cognitive implications (Maudsley, 2002). Other work, however, did not provide information about the theoretical approaches adopted (Premi et al., 1994; Wilkerson & Maxwell, 1988). As Newton et al (2001) indicate "*the theory is a bit vague*" (p.222). This restricts credence in the research field (Rudestam & Newton, 1982). Some studies did contextualise the rationale for research from the existing empirical evidence (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Vernon, 1995). As Colliver (2000), however, argues "*the ties between educational theory and research (both basic and applied) are loose at best*" (p.259).

Research from Bernstein et al (1995) and Kaufman and Holmes (1996) had clearly based research questions generated from the literature. These were

delineated in relation to distinctive aspects for students and staff in examining the tutoring process. Other studies, however, lacked research questions and hypotheses reducing usability in practice (Carnine, 1997). These studies also failed to detail the relationships between variables and the espoused direction of causality (Houlden et al., 2001; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988.) In this sense the majority of work lacked the level of scientific rigour to explicitly define hypotheses and variables in testable form which undermines credibility (Cameron, 2004; Friedman et al., 1990; Greenhalgh, 2001.)

3. Research Design

The majority of research used semi-structured questionnaires to examine shifts in attitudes of key players in both the short term (Bernstein et al., 1995), longer term (Kaufman & Holmes, 1996) and at a snapshot (Vernon, 1995). Such techniques are appropriate to the field of study given the size of samples involved. None of these studies, however, provided a rationale to justify the use of a questionnaire over other methods, nor the theoretical foundations underpinning a qualitative approach. It is not until later research that such decisions are justified with reference to the methodology adopted (Houlden et al., 2001; Maudsley, 2002). These pieces also extend the qualitative approach by complementing questionnaires with semi-structured interviews to illuminate findings and check on accuracy. Overall, however, a substantial sector of studies fails to provide sufficient information as to the design which makes replication difficult.

Furthermore the papers do not "*in general appear to contain sufficient description of either the experimental or control interventions*" (Newman, 2004) (p.8). In addition randomized designs were not utilized despite the fact that they "*deliver the most trustworthy data*" (Dochy et al., 2003) (p.541). In fact studies were of limited reliability as randomisation did not occur nor was it compensated for by controlling or matching participants along key factors (Colliver, 2000; Dochy et al., 2003; Smits et al., 2002). Newman et al (2001), likewise, highlight methodological weaknesses of "*single group post-test designs, the use of non comparable sample groups, poorly specified and or non-comparable interventions*" (p.222). Ethical considerations connected to the sample selected are similarly scant with only two pieces scrutinizing these in depth (Kaufman & Holmes, 1996; Maudsley, 2002). This casts doubt as to the applicability of studies to students and staff alike (Greenhalgh, 2001).

No study adequately discussed or controlled for the potential range of confounding variables in the research design. For example Newman (2004) highlights how "*study design, randomisation, level of education and assessment format are all potential moderating variables*" (p.7) yet no paper addressed these. Many of the studies ignored the role of researcher bias in the collection and interpretation of qualitative data (Bernstein et al., 1995; Kaufman & Holmes, 1996; Houlden et al., 2001; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988). Indeed in all work the researcher was not blind to group allocation and was actively involved in programme delivery and data analysis, which introduces subjectivity into

proceedings (Greenhalgh, 2001). It is recognised, however, that the practicality of such blind allocation in educational interventions is problematic. In addition no procedures for researcher reflexivity to modify the impact of bias are detailed which undermines the scientific rigour of studies (Sherrard, 1997). Variables are only provided with clear category choices and defensible cut off points in two papers in terms of student and staff samples and questionnaire outcomes and scoring (Kaufman & Holmes, 1996; Vernon, 1995). As Colliver (2000) points out variables are "*lacking explicit descriptions of their interrelationships and of their relationship with observables*" (p.264). This restricts replication (Greenhalgh, 2001; Rudestam & Newton, 1982).

The measurement devices used in scientific research need to be justified for study (Connor, 1997). In some papers the selection of semi-structured interviews to extend the survey was explained in ensuring findings were ecologically more valid (Maudsley, 2002; Wilkerson & Maxwell, 1988). No other study provided any rationale for measures utilised (Bernstein et al., 1995; Houlden et al., 2001; Maudsley, 2002; Premi et al., 1994; Vernon, 1995). Furthermore no external validation occurred for any outcome measure (Friedman et al., 1990). Technical detail was reported on in Vernon's 1995 study alone whilst other pieces failed to comment on the scientific rigour of measures used. As Newman et al (2001) demonstrate, PBL studies generally, tend to "*use non-objective outcome measures/assessment tools that had no established validity and or reliability*" (p.222). Only one study discussed the piloting procedures and question selection of a device

(Maudsley, 2002). Moreover, the procedures for administration were limited, with papers neglecting to examine the role of the researcher in the administration and interpretation of measures (Bernstein et al., 1995; Kaufman & Holmes, 1996; Houlden et al., 2001; Premi et al., 1994; Vernon, 1995). These methodological weaknesses limit research in this domain.

All studies examined samples of staff and students which were appropriate to the matters being explored. In fact as Greenhalgh (2001) indicates such research in real life contexts provides ecological validity. Certain work built on this by including demographic characteristics of the sample selected (Kaufman & Holmes, 1996; Maudsley, 2002; Vernon, 1995; Wilkerson & Maxwell, 1988). This allows for extrapolation to the population and enhances generalisation. Only one study, however, explored the recruitment and selection of sample including statistics of response rates and exclusion (Vernon, 1995). In this manner other papers were restricted with issues of self-selection left unaddressed. This may have introduced bias into the research as subjects amenable to PBL were potentially used in that only tutors who felt positively about PBL responded (Colliver, 2000; Greenhalgh, 2001). Also some studies used only tutors who were interested in PBL as opposed to the whole of a faculty which may have positively skewed results (Vernon, 1995; Wilkerson & Maxwell, 1988). In addition some studies used only a snapshot of the sample from a single academic year which limits usability (Bernstein et al., 1995; Houlden et al., 2001; Maudsley, 2002; Premi et al., 1994; Vernon, 1995). Finally no study reported effect size despite the fact that "*sample size and statistical power emerge as important*

considerations" (Friedman et al., 1990) (p.13). Such an omission limits credence as does the lack of information on differences at baseline (Colliver, 2000; Dochy et al., 2003; Greenhalgh, 2001).

4. Results and Discussion

The adequacy of results varied considerably in the studies scrutinized. Some pieces described statistical techniques employed in analysing quantatitive data from questionnaires (Bernstein et al., 1995; Kaufman & Holmes, 1996; Vernon, 1995) whilst others omitted such information. The more qualitative research often lacked any detail on the method of interpretation, which restricts credibility (Bernstein et al., 1995; Kaufman & Holmes, 1996; Premi et al., 1994; Wilkerson & Maxwell, 1988). Two studies based qualitative analysis on the principles of grounded theory (Houlden et al., 2001; Maudsley, 2002). These studies rooted the analysis in an accountable manner to qualitative research methods unlike the other papers. The majority of work, likewise, failed to discuss the range of confounding variables inherent in design. In fact the direct role of the researchers was ignored with no reflexivity checks with participants (Stevenson & Cooper, 1997). The only studies to account for possible confounds were Kaufman and Holmes (1996) who discussed the novelty value of PBL and Vernon (1995) who discussed the self-selection of samples. Consequently most pieces failed to accommodate for the potential range of confounding factors in data analysis from the role of the educational environment to the influence of differing durations of exposure to PBL practice.

In terms of conclusions drawn, some are consistent with a systematic use of quantitative and qualitative data. Moreover these pieces contextualise the findings in the wider research field (Bernstein et al., 1995; Kaufman & Holmes, 1996; Maudlsey, 2002; Vernon, 1995). Papers were all limited, however, in the over generalisation of positive results with no raw questionnaire or interview data with which to compare results (Erlandson et al., 1993). Many studies also failed to consider alternative conclusions consistent with the data which may mean findings are skewed (Bernstein et al., 1995; Kaufman & Holmes, 1996; Houlden et al., 2001; Maudsley, 2002; Premi et al., 1994; Wilkerson & Maxwell, 1988). Few evaluations address limitations with only Maudsley (2002) critiquing the telephone interview technique and Vernon (1995) the confounds of the survey and variability in PBL programs. Practical implications are, however, considered across all pieces especially the need to support tutors in the transition to PBL (Bernstein et al., 1995; Kaufman & Holmes, 1996; Houlden et al., 2001; Maudsley, 2002; Premi et al., 1994; Wilkerson & Maxwell, 1988). In this fashion the research has high ecological validity to numerous universities in the management of change (Friedman et al., 1990). Theoretical considerations are not discussed in any piece indicating an area of further development as Colliver (2000) states "*differences would require explanation in terms of educational principles and underlying theoretical mechanisms*" (p.263). Furthermore future research is only detailed in a few pieces (Premi et al., 1994; Wilkerson & Maxwell, 1988) which restricts the evolution of the evidence base.

5. Summary of Psychological Literature

All studies reviewed effectively examined the key stakeholders in response to curricula change and the impact on tutor (Friedman et al., 1990). Early work from Premi et al (1994), Bernstein et al (1995) and Kaufman and Holmes (1996) offered an insight into students' and staffs' shifting attitudes. The latter two studies were of particular practical utility in suggesting implementation procedures for future use. Subsequent work focused on tutors in particular with Vernon (1995) providing an extensive study of tutors from over twenty two universities. Houlden et al (2001) and Maudsley (2002) outlined tutor perceptions in terms of the problems and processes to identify practice implications. Considerable issues exist, however, within studies as to experimental design as Newman (2004) argues papers "*do not provide high quality evidence with which to provide robust answers*" (p.5).

A number of key themes recur throughout the literature listed. These are discussed and then followed up in full in section 4. Firstly an area of consistency was the *positive way in which PBL was viewed* by students and staff alike (Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Vernon, 1995; Premi et al., 1993). In particular it was demonstrated that direct exposure and experience with PBL "*led to more favourable attitudes among the students and faculty*" (Bernstein et al., 1995) (p.245). Such findings were shown to be substantiated over time (Maudsley, 2002). The work also evidenced that PBL was evaluated more positively than traditional techniques on a number of aspects from tutors' personal satisfaction to levels of student interest (Houlden et al., 2001; Kaufman & Holmes, 1996;

Maudsley, 2002; Premi et al., 1993; Vernon, 1995). Additionally an area of agreement was the *critical importance of the tutor in the PBL process* (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1993; Vernon, 1995; Wilkerson & Maxwell, 1988). As Mayo et al (1995) argue "*the PBL tutors set the stage for learning and present themselves as models of the learning process*" (p.126).

An area of controversy prevalent in the research pertains to the *impact of the approach on students' knowledge acquisition* as perceived by tutors. Some studies demonstrated that tutors preferred traditional teaching techniques over PBL in terms of learners' basic knowledge development (Kaufman & Holmes, 1996; Vernon, 1995). As Kaufman and Holmes stated there were "*a number of staff who believe that students' factual knowledge has declined" (*p.374). In contrast other work highlighted the advantages of PBL as "*it teaches students how to learn rather than to memorise*" (Bernstein et al., 1995) (p.245). Other variability across studies relates to *PBL processes*; from changing tutorial groups and team tutoring (Kaufman & Holmes, 1996) to sustained PBL groups across the year with the same tutor (Bernstein et al., 1995; Maudsley, 2002). Concerns were also raised surrounding the *"specific steps (the method) within PBL sessions*" (Bernstein et al 1995) (p.245). These issues will need addressing to detail the tutor's role in effective PBL delivery.

A key practice question emerging from the literature concentrates on the *tutor's remit and responsibilities in PBL sessions*. The importance of the tutor

is perceived as paramount across studies (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1993; Vernon, 1995; Wilkerson & Maxwell, 1988). As Mayo et al (1995) describe from an experiential perspective, having implemented the approach; "*PBL sessions reflect the tutors imagination, creativity, personality, and temperament. These sessions succeed or fail in direct proportion to the tutors preparedness and training for the task*" (p.126). Papers reviewed revealed important issues surrounding group facilitation skills of tutors in terms of the level of intervention and knowledge provided (Kaufman & Holmes, 1996; Maudsley, 2002; Wilkerson & Maxwell, 1988). As Kaufman and Holmes (1996) detail "*faculty tutors were unsure about when to intervene either to guide the group, correct misinformation, or explain case material*" (p.374). It appears the balance tutors must strike between leading a group and withdrawing from it may need closer examination (Bernstein et al., 1995; Kaufman & Holmes, 1996; Maudsley, 2002; Vernon, 1995; Wilkerson & Maxwell, 1988). These matters have been evidenced in recent research within educational psychology training in Dundee which identified tutor role ambiguity and facilitation skills as issues (Kerr & Smith, 2005). As Maudsley (2002) argues "*more empirical evidence about the tutors influence might help to adapt PBL (and the tutors role) appropriate to students progression* "(p.171). This could be established through research to determine tutor impact on the group/individual in PBL sessions from a content and process perspective.

An additional but linked unanswered practice question relates to *tutor performance* in PBL and its impact on promotion and recruitment (Kaufman & Holmes, 1996; Wilkerson & Maxwell, 1988). The end of unit evaluation in the PBL process provides a means of judging teaching effort and quality. As Kaufman and Holmes (1996) suggest such information could assist in tenure decision making in a systematic manner. Similarly Wilkerson and Maxwell (1988) advocate it as a "*way of producing evidence of teaching effectiveness to be used for promotion*" (p.898) as well as a motive for recruitment with "*the opportunity to collaborate with faculty members*" (p.897).

The majority of studies demonstrated *unresolved issues in connection to the training in PBL*, in particular with a need for an induction programme (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002). Indeed as Houlden et al (2001) specified " *We believe that the problems identified by the tutors in our survey would be overcome if more instruction in PBL processes and skills were provided for both students and facilitators*" (p.81). More specifically the teaching staff identified important areas for development as "*the need for further training in intervening appropriately, dealing with difficult situations, questioning techniques and evaluating students*" (Kaufman & Holmes, 1996) (p.376). The *assessment issue* was also prevalent throughout papers with questions raised as to the best method by which to examine student's knowledge and skills development (Kaufman & Holmes, 1996; Maudsley, 2002; Vernon, 1995). In fact as Newman (2004) pinpoints "*there is no consensus on either the*

outcomes or methods of measurement that should be used to assess the effects of PBL" (p.23).

Research into the role of tutor in PBL has principally used qualitative methodologies on student and staff samples (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988). According to Friedman et al (1990) there is a "*difficulty of conducting evaluation studies that will be credible in the scientific community" (*p.9). As Smits et al (2002) comment of PBL research generally *"there have been few well conducted trials*" (p.153). In the literature reviewed the size of samples has been small despite the fact that "*sample sizes must be sufficient to detect the differences that are educationally meaningful*" (Friedman et al., 1990) (p.13). Furthermore the research has not provided information on the reflection and metacognitive aspects of PBL, which is specifically encouraged in the model (Colliver, 2000; Newman, 2004). Finally the potential benefits to life-long learning and continuing professional development for students and staff alike have not been addressed. All such areas are potential avenues for future research.

Section 4: Integration of Theory, Research and Practice

This section outlines the way in which current research, theory and practice overlap and contradict in the professional training of practitioners using PBL with regard to the role of tutor. Initially areas of agreement and conflict will be examined followed by a focus on how these can be addressed in both the short, medium and long term.

Research and Practice Issues

The need for an alternative approach to the training of professionals is evident both at the university level and at a societal level in response to market forces (Dochy et al., 2003; Fraser & Greenhalgh, 2001; Henlow & Evensen, 2000; Macdonald & Savin-Baden, 2004). As Dochy et al (2003) highlight "*an important challenge for today's higher education is the development and implementation of instructional practices that will foster in students the skills to apply knowledge effectively*" (p.523). Such a motive is paralleled in research with tutors keen to investigate innovative curricula designs to meet such demands (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988). Agreement exists across research and practice in the need for curriculum evolution within a changing theory and community context. Furthermore the role of tutor is considered critical in the implementation of PBL change both theoretically (Barrows, 1986; Hmelo-Silver, 2004; Macdonald & Savin-Baden, 2004) and empirically (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988). In connection to this the need for tutor transformation is recognised (Barrows, 1986; Hmelo-Silver, 2004; Macdonald & Savin-Baden, 2004) as is the need for training in the transition of tutors to PBL facilitators (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988).

There is general agreement in research and practice as to the changing of tutor role in PBL. Indeed the shifts in skills are encapsulated as Hmelo–Silver (2004) details "the *teacher acts to facilitate the learning process rather than to promote knowledge*" (p.235). Despite this, however, controversy exists as to both the acceptance of such changes on behalf of tutors and the processes pertaining to PBL. In particular some work suggests tutor adjustment difficulties (De Weerdt et al., 2002; Houlden et al., 2001; Tohey, 1999; Ryan, 1997; Vernon, 1995) whilst other papers indicate positive tutor acceptance (Bernstein et al., 1995; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Wilkerson & Maxwell, 1988). This represents an area of further analysis, as Savin–Baden (2004) indicates "*for staff, the challenges of using problem based learning are equally complex in that they relate not only to issues of teaching and learning; but also to the personal challenges that emerge as students question their perspectives and prior experience*" (p.1/2).

In this fashion university practice and research must endeavour to explore areas of disparity to devise induction and support packages to implement PBL smoothly (Bernstein et al., 1995; Kaufman & Holmes, 1996; Maudsley, 2002). Moreover discrepancies between research and practice pertaining to the PBL model of use (Vernon, 1995) or changing of tutor groups (Kaufman & Holmes, 1996) could inform policy decisions. Work between and across universities could address the steps (Maudsley, 2002) and structure of PBL (Bernstein et al., 1995) and consider translation to LEA practice settings.

Short Term Action to Address Practice Questions

A review of both research and practice indicated a number of unanswered practice questions pertaining to the tutor's role in PBL. The key issues were how to facilitate; how to model the learning process and at the same time teach the content (Bernstein et al., 1995; Kaufman & Holmes, 1996; Maudsley, 2002; Newman, 2004; Wilkerson & Maxwell, 1988). Short term action at the LEA and HEI policy level could jointly devise clear guidelines for each respective remit. Firstly models of teaching and adult learning could be specified from the Cognitive Apprenticeship Model (Collins et al., 1989) to critically responsive teaching techniques (Brookfield, 1987). As Zinkiewicz et al (2003) propose these could be promoted in the context of group development theories (Tuckman & Jensen, 1977) and group cohesiveness (Michaelsen et al., 1997). Consultation with tutors and supervisors regarding the use of these could translate theory from policy to practice. Indeed the relevance of PBL given the need for multi-professional team work in the LEA context could be emphasised. This would ensure the salience of PBL for EP practice in light of legislative shifts towards team based multi-agency work (DfES, 2004).

Facilitation skills are, as Kaufman and Holmes (1996) detail, "*an area to be reassessed*" (p.374). At a policy level specific strategies for facilitatory teaching could be developed drawing on metacognitive questioning (Hmelo-Silver, 2002); alongside modelling and coaching (Hmelo-Silver & Barrows, 2002) to scaffold students' learning in a group context (Hmelo-Silver, 2004). This meets the need highlighted in research as "*tutors wished to receive*

further training in questioning techniques. In particular their ability to interject higher level questions without disrupting or derailing the group process" (Kaufman & Holmes, 1996) (p.374). Furthermore the skills of modelling the process of learning could be delineated using both wandering facilitation and student facilitation (Hmelo-Silver, 2004). As Newman et al (2001) state "*assistance for learning is provided through interactions characterised by such activities as directing, modelling, questioning and providing cognitive structuring and feedback*" (p.221).

The use of questioning, scaffolding and promoting reflection through feedback all need integrating to support cumulative learning (Exley & Dennick, 2004). Such skills would need consultation with tutors and supervisors by which to clarify methods of practice. Moreover these would need to be enshrined in working agreements and arrangements across LEA and HEI contexts for consistency with workshops and feedback mechanisms (Barrows, 2000). Joint training of supervisors and course tutors could occur in the techniques and types of questioning to support reflection and scaffold EPiTs progression through PBL and practice. This would not only support the continuing professional development of those new to the discipline but established practitioners also. In this manner deep level learning and reflection could potentially occur for all involved.

A final unanswered practice question related to the transmission of content knowledge in PBL sessions from the tutor. As Maudsley (1999) indicates " *the tutors challenge is to forego the gratification of dispensing facts, and walk*

the tightrope of effectiveness by balancing intervention in the group process between an informal, empathetic style and subtle sparing use of personal context expertise" (p.661). In future accurate evaluation will be needed in the university and LEA context to consider content expertise and process facilitation skills of tutors and supervisors in PBL scenarios (Newman, 2004; Maudsley, 1999). As Barrows (1986) argues "*the quality of tutorial skills is a common concern of schools that use problem based approaches*" (p.485). This provides a further motive for careful assessment through tutor and student feedback of skills usage. This could also inform tutor promotion and recruitment, issues raised in the research (Kaufman & Holmes, 1996; Maudsley, 1999; Wilkerson & Maxwell, 1988). Furthermore such evaluations could act as a quality control device for both students and staff alike. In this way discrepancies and differences between different facilitator styles could be examined which, in turn could influence tutor training and development.

Medium Term Action to Resolve Identified Issues

A key unanswered issue in the literature was the need for tutor training in the methods of PBL (Bernstein et al., 1995; Houlden et al., 2001; Kaufman & Holmes, 1996; Maudsley, 2002; Premi et al., 1994; Vernon, 1995; Wilkerson & Maxwell, 1988). Newman's 2004 overview indicated this also as an essential and yet often omitted aspect of PBL implementation. This is despite its importance as recognised by curriculum innovators (Evans & Taylor, 1996; Holmes & Kaufman, 1994; Hitchcock & Mylona, 2000; Wetzel, 1996). As Evans and Taylor (1996) argue *"in the context of major curriculum change a staff development programme has a crucial role to play* "(p.366) being "*the*

most essential ingredient in the success of any educational innovation" (Wetzel, 1996) (p.474). Given its critical relevance medium term action must design, implement and evaluate a tutor/supervisor development process to ensure successful curricula change (Holmes & Kaufman, 1994; Irby, 1996). Cross course collaborations between universities could occur to investigate and inform training procedures. In particular a working party across HEI and LEA domains could support evolving PBL practice.

Initially tutors need to feel informed and involved in the changing curriculum alongside contributing to its evolution (Engel, 1998). A variety of models could then be drawn upon to develop an induction/support package to develop staff skills specifically. 4 key strategies based on Fullan and Pomfret's 1977 framework could underpin such a development programme. This proposes close integration of in-service training, resource support, feedback mechanisms and participation in decision making. Initial training could provide an induction to the processes of PBL and the role of tutor (Evans & Taylor, 1996; Holmes & Kaufman, 1994; Wetzel, 1996). This could include day long workshops to present the principles of PBL followed by observation of a simulated live role play of a PBL session (Holmes & Kaufman, 1994; Irby, 1996; Wilkerson & Maxwell, 1988). This could then lead into a "theory in practice" part whereby tutors experiment with the techniques (Barrows, 1988) with groups of volunteer students (Holmes & Kaufman, 1994) or with colleagues on a meta-problem (Evans & Taylor, 1996). As Evans and Taylor (1996) argue "*presenting PBL in an experimental format rather than didactic, was found to be a successful way of introducing staff to*

this style of learning" (p.366). Tutors could receive feedback and pool together knowledge and skills to provide future resource support following this (Holmes & Kaufman, 1994; 1996; Fullan & Pomfret, 1977; Wetzel, 1996).

The support of tutors could continue using the 4 stage tutor training programme of Harvard (Wetzel, 1996) or the 7 stage model from Dalhousie (Holmes & Kaufman, 1994). Both begin with *initial orientation workshops* (as outlined) followed by *unit orientation and refresher meetings*. These provide previews and discussions surrounding the units to ensure content clarity both 6-8 weeks prior to the unit and one week before implementation. *Weekly tutor meetings* could also be included to provide a feedback function to discuss issues arising and evaluate cases completed. Such a mechanism would allow colleagues to reflect on the process and content alongside participating in decision making (Fullan & Pomfret, 1977; Irby, 1996; Wilkerson & Maxwell, 1988). *Tutorial observations*, likewise, could form part of the programme using peer observation networks to inform skills development (Holmes & Kaufman, 1994; Wetzel, 1996; Wilkerson & Maxwell, 1988).

Evaluations of tutors could also be made by students with end of unit questionnaires (Holmes & Kaufman, 1994; Wilkerson & Maxwell, 1988). Whilst *Unit evaluations* could be completed by tutors at the end of a unit to review success and discuss future improvements. As Wilkerson and Maxwell (1988) highlight this "*creates opportunities for educators and faculty members to talk together about the new curriculum*" (p.898). This involves

tutors in the evolution of the curriculum alongside their own continuing professional development (Holmes & Kaufman, 1994; 1996). Such evaluations and input are needed in the realms of research as "*few studies document the impact of PBL faculty development strategies*" (Hitchcock & Mylona, 2000) (p.52). Student orientation and induction also warrant input and investigation (Wetzel, 1996).

Faculty development could alternatively be based on other models in the research drawing on either general skills, developmental, comprehensive, or course based models (Irby, 1996; Wilkerson & Irby, 1998). As Irby (1996) argues all such frameworks incorporate general and course specific instructional level development but vary according to leadership and organisational evolution. In all such practices three critical areas of tutor training are needed namely "*general tutor skills, content-specific tutor knowledge and skills, and advanced knowledge and skills*" (Irby, 1996) (p.79). All such domains will require development in HEI and LEA contexts. This could occur on a national level with cross university and placement collaborations aimed to share productive PBL practice.

Another unresolved issue identified in the research (Kaufman & Holmes, 1996; Maudsley, 2002; Vernon, 1995) and practice (Elton, 2004; Hmelo-Silver, 2004; Newman, 2004; Macdonald & Savin-Baden, 2004) were the difficulties tutors found in assessing students. As Savin-Baden (2004) illustrate "*assessment currently appears to be one of the most controversial concerns in problem based learning*" (p.223). Medium term action at the level

of university can address this to devise "assessments *that enable students to grow and develop*" (Savin-Baden, 2004) (p.224). Basic principles could underpin development with reliability, validity, fairness, and standardisation adhered to (Newstead, 2002 in Zinkiewicz et al., 2003). In addition as Smits et al (2002) argue "*outcome variables should correspond with our educational objectives* "(p.155). Macdonald and Savin-Baden (2004) and Biggs (1999) reiterate this, detailing how an alignment between learner outcomes, teaching methods and assessment should be utilised.

Varied assessment techniques could be trialled to encompass learning approaches, content knowledge and application to practice (Newman, 2004). Assessments could "*be based in practice contexts in which students will find themselves in the future- whether real or simulated*" (Macdonald & Savin-Baden, 2004) (p.7). This would inform assessment as to the application rather than reproduction of knowledge (Dochy et al., 2003) and has been used in written tasks (Gijbels, et al., 2005). Present PBL developments in Dundee on the MSc training course in educational psychology have revised assessment techniques to incorporate more formative assessments to offer a better balance between product and process (Kerr & Smith, 2005). Web based approaches such as Virtual Learning Environments (VLEs) could also be explored.

Approaches could also "*help students to become stakeholders in the assessment process* "(Savin-Baden, 2004) (p.232). This could include peer and self assessment in line with the principles of self-directed learning

intrinsic to PBL (Dochy et al., 2003; Macdonald & Savin-Baden, 2004; Savin-Baden, 2004). Universities will need to develop assessment approaches alongside students and LEA colleagues that fit both with PBL and the professional practice requirements in context. Working groups across HEIs and LEAs could support such innovations (Schwartz, 2004).

Long Term Action to Explore Future Research Directions

An examination of both research and practice revealed a range of opportunities for future development. As numerous critics claim the literature lacks the scientific rigour of high quality research being predominately of qualitative design with restricted outcomes (Colliver, 2000; Newman, 2004; Norman et al., 2005). In future designs could examine outcomes more explicitly (Chen et al., 2004; Friedman et al., 1990). Newman et al., (2001) suggest the use of a range of measures to tap into group processes, group cohesiveness and group problem solving. Similarly the SOLO taxonomy (Structure of Observed Learning Outcomes) could be used to gauge student's individual development (Biggs, 1999). Evidence exploring the endurance of effects and transfer to alternative contexts could be explored (Friedman et al., 1990; Norman et al., 2005). An example of this is the longitudinal study conducted in Dundee to obtain information on the longer term impact of PBL on professional EP practice (Kerr & Smith, 2005). Furthermore studies could become more theoretically driven (Friedman et al., 1990; Norman & Schmidt, 2000) alongside using comparative designs (Distlehorst et al., 2005; Newman, 2004). Finally the use of larger sample

sizes within more rigorous research designs would provide greater credibility to research (Friedman et al., 1990; Greenhalgh, 2001).

The scope of PBL research could be extended to determine what skills make a tutor effective using structural-equation modelling (Norman & Schmidt, 2000; Schmidt & Moust, 1995). Similarly characteristics of the ideal PBL tutor could be explored through student perceptions (Lin, 2005; Mayo et al., 1995) or the Tutor Intervention Profile (TIP) (De Grave et al., 1999). The possibilities of computer based learning to support PBL approaches could also be investigated (Cook, 2005; Koschmann et al., 1996). The use of computer mediated conferencing to support educational psychologists in training across HEI and LEA contexts has proved fruitful (Boyle et al., 2003). In this manner web based learning and virtual learning environments could be examined as they emulate the reflection and transactive discussion inherent in PBL (Cook, 2005; Koschmann et al., 1996; Zinkiewicz, et al., 2003). Work could also examine the impact of PBL on learning as "*more research of self-directed learning is sorely needed*" (Colliver, 2000) (p.266). Learning styles could be evaluated using approaches such as ASSIST (Approaches and Study Skills Inventory; Tait & Entwistle, 1996). Finally the tutor/supervisor role in implementing PBL could be explored as a means of continuing professional development (Stefani & Elton, 2002).

Concluding Comments

"*PBL represents a major development and change in educational practice that continues to have a large impact across subjects and disciplines worldwide*" (Newman, 2004) (p.5). As this paper demonstrates such an approach can meet societal demands and the continuing evolution of the discipline. The literature review mirrors the need and strengths of PBL with tutors as key to effectiveness. Future work is necessitated to prepare both HEIs and LEAs in the provision of programmes to translate PBL principles into practice. This heralds an innovative time for EP training, the potential benefits of which will be realised throughout the profession.

References

Albanese, M.A., and Mitchell, S. (1993). Problem-based learning: A review of the literature on its outcomes and implementation issues. Academic Medicine, 68, (1), 52-81.

Argyris. C, and Schon, D. A. (1974). Theory in Practice. San Francisco: Jossey-Bass.

Barrows, H.S. (1986). A taxonomy of problem based learning methods. Medical Education, 20, 481 – 486.

Barrows, H.S. (1994). Problem-based learning applied to medical education. Springfield, Illinois: Southern Illinois University School of Medicine.

Barrows, H.S. (1988). The Tutorial Process. Southern Illinois University School of Medicine, Springfield, Illinois. pp.43-45.

Barrows, H.S. (2000). Forward. In D. H. Evensen, & C. E Henlow (Eds.), Problem based learning: A research perspective on learning interactions (pp.vii-ix). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Berkson, L. (1993). Problem-based learning: Have the expectations been met? Academic Medicine. 68, (10), 79-88.

Bernstein, P., Tipping, K., Bercovitz, K. and Skinner, H.A. (1995). Shifting student and faculty to a PBL curriculum: attitudes changed and lessons learned. Academic Medicine, 70, 245-247.

Biggs, J. (1999). Teaching for Quality Learning at University. Buckingham: The Society for Research into Higher Education / Open University Press.

Boud, D.J. (ed.) (1985). Problem-Based Learning in education for the professions. Sidney: Higher Education Research and Development Society of Australia.

Boud, D., and Feletti, G. (Eds.) (1997). The challenge of problem based learning (2nd Ed). London: Kogan Page.

Boyle , J., Tolnie, A., and Stobie, I. (2003). Bridging the Gap: probationer educational psychologists' use of computer-mediated conferencing. Educational Psychology in Practice, 19, (3), 189-197.

British Psychological Society. (BPS). (2004). Report from the Working Party on Educational Psychology Training in England, Wales and Northern Ireland. British Psychological Society.

Brookfield, S. D. (1987). Developing Critical Thinkers: Challenging Adults to explore Alternative Ways of Thinking and Acting. San Francisco, CA: Jossey-Bass.

Broughton, J. M. (1981). Piaget's structural developmental psychology: Piaget and Structuralism. Human Development, 24, 78-109.

Cameron, R. J. (2004). Critiquing a research paper: suggested guidelines. University College London. A copy is available from University College London, Educational Psychology Group.

Carnine,D. (1997). Bridging the research to practice gap. Exceptional Children. 63 (4), 513-521.

Chen, F. M., Bauchner, H., and Burstin, H. (2004). A Call for Outcomes research in Medical Education. Academic Medicine, 79, 955-960.

Cline, T., Frederickson, N., and Wright, A. (1990). Effective In-service Training: A Resource Pack for LEA INSET Providers. London: Educational Psychology Publishing.

Collins, A., Brown, J. S., and Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In Resnick, L. B. (ed). Knowing, Learning, and Instruction: Essays in Honour of Robert Glaser. Erlbaum, Hillsdale, NJ, pp. 453-494.

Colliver, J.A. (2000). Effectiveness of Problem-based Learning Curricula: Research and Theory. Academic Medicine, 75, (3), 259-266.

Connor, M.J. (1997). Critical Appraisal. An unpublished paper available from Surrey Educational Psychology Service.

Cook, D. A. (2005). The Research We Still Are Not Doing: An Agenda for the Study of Computer-Based Learning. Academic Medicine, 80, 541-548.

De Grave, W. S., Dolmans, D., and Van-der-Vleuten, C. (1999). Profiles of Effective Tutors in Problem-based learning: Scaffolding student learning. Medical Education, 33, (12), 901-906.

De Weerdt, S.D. Corthouts., F., Martens,H. and Bouwen, R. (2002). Developing Professional Learning Environments: models and application. Studies in Continuing Education, 24, (1), 25- 38.

Department for Education and Skills (DfES). (2004). Every Child Matters: Working with voluntary and community organisations to deliver change for children and young people. Nottingham: DfES Publications.

Distlehorst., L. H., Dawson, E., Robbs, R.S. and Barrows, H.S. (2005). Problem-Based Learning Outcomes: The Glass Half-Full. Academic Medicine, 80, 294-299.

Dochy, F., Segers, M., Van den Bossche, P. and Gijbels, D. (2003). Affects of problem based learning: a meta-analysis. Learning and Instruction, 13, 533– 568.

Dundee Educational Psychologists in Training. (2005). Emotions within Problem Based Learning. Paper presented at the National EPiT Conference, Bristol, 6th- 8th July 2005.

Educational Psychologists in Training (EPiT) National Conference. (2005). Bristol, 6th-8th July 2005.

Elton, L. (2004). Problem Based Learning for Beginners. Presentation at University College London, June 2004.

Engel, C.E. (1998). Not Just a Method But a Way of Learning. In Boud, D., and Feletti, G. (Eds.) (1997). The challenge of problem based learning (2nd ed.). London: Kogan Page.

Erlandson, D. A.,Harris , E.L. Skipper, B.L. and Allen, S.D. (1993). Doing Naturalistic Inquiry. London: Sage.

Evans, P. A. and Taylor., D.C.M. (1996). Staff Development of tutor skills for problem-based learning. Medical Education, 30, 365-366.

Exley, K., and Dennick. R. (2004). Small group teaching, tutorials, seminars and beyond. London: Routledge Falmer.

Fraser, S.W. and Greenhalgh, T. (2001). Coping with complexity: educating for capability. British Medical Journal, 323, 799-803.

Frederickson, N. (1993). Using Soft Systems Methodology to rethink special educational needs. In A. Dyson, and C. Gains (Eds.), Rethinking Special Needs in Mainstream Schools: Towards the Year 2000 (pp.1-21). London: David Fulton.

Frederickson, N., Reed, P., Savage, R., Murphy, R., and Turner, M. (2003, September). Research Methods and Statistics for Educational Psychologists. http://www.webct.ucl.ac.uk:8900/

Friedman, C. P., De Bliek., R., Greer. D.S., Mennin, S.P., Norman., G.R., Sheps., C. G., Swanson, D. B and Woodward, C.A. (1990).Charting the winds of change: Evaluating innovative medical curricula. Academic Medicine, 65, 8-14.

Fullan, M., and Pomfret, A. (1977). Research on Curriculum and instruction implementation. Review of Educational Research, 47, 335-397.

Gijbels, D., Van-de-Watering, G. and Dochy, F. (2005). Integrating assessment tasks in a problem-based learning environment. Assessment and Evaluation in Higher Education, 30, (1), 73-86.

Greenhalgh, T. (2001). How to read a paper. The basics of evidence based medicine. BMJ Books: London.

Henlow, C. E., and Evensen, D. H. (2000). Introduction to Problem Based Learning: Gaining Insights on Learning Interactions through Multiple Methods of Enquiry. In D. H. Evensen, and C. E Henlow (Eds.), Problem based learning: A research perspective on learning interactions (p. 1-16). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Hitchcock, M. A. and Mylona, Z. H. E. (2000). Teaching faculty to conduct problem-based learning. Teaching and Learning in Medicine, 12, (1), 52-57.

Hmelo-Silver, C. E. (2002). Knowledge recycling; Crisscrossing the landscape of educational psychology in a Problem-Based Learning Course for pre-Service Teachers. Journal of Excellence in College Teaching, 11, 41-56.

Hmelo-Silver, C. E. (2004). Problem Based Learning: What and How Do Students learn? Educational Psychology Review, 16, (3), 235- 266.

Hmelo-Silver, C. E., and Barrows, H.S. (2002). Goals and strategies of a constructivist teacher. Paper presented at American Educational Research Association Annual Meeting, New Orleans, LA.

Hmelo, C.E., and Lynn, X. (2000). Becoming self directed learners: strategy development in problem based learning. In D. H. Evensen, & C. E Henlow (Eds.), Problem based learning: A research perspective on learning interactions (pp. 1-16). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Holmes, D. B. and Kaufman. D. M. (1994). Tutoring in problem-based learning: a teacher development process. Medical Education, 28, 275- 283.

Houlden, R.L., Collier. C.P., Frid, P. J. John, S. L., and Pross, H. (2001). Problems Identified by Tutors in a Hybrid Problem-base Learning Curriculum. Academic Medicine, 76, (1), 81.

Irby, D. M. (1996). Models of Faculty Development for Problem-Based Learning. Advances in Health Sciences Education, 1, 69-81.

Kaufman, D. M. and Holmes, D. B. (1996). Tutoring in problem-based learning: perceptions of teachers and students. Medical Education, 30, 371-377.

Kaufman, D. M. and Holmes, D. B. (1998). The relationship of tutors' content expertise to interventions and perceptions in a PBL medical curriculum. Medical Education, 32, (3), 255-261.

Kerr, C. and Smith, E. F. (2005). An evaluation of the perceived effectiveness of problem-based learning on the Dundee MSc Educational Psychology training programme. E-journal@DundeeUniversity.co.uk.

Kolb, D. A. (1984). Experiential learning: Experience as the source of learning and development. Englewood Cliffs, NJ: Prentice-Hall.

Koschmann, T.D., Kelson, A.C., Feltervich, P.J., and Barrow, H.S. (1996). Computer supported problem based learning: a principled approach to the use of computers in collaborative learning. In T.D. Koschmann (Ed.) CSCL: Theory and Practice of an Emerging Paradigm. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Lin, C. S. (2005). Medical students' Perception of Good PBL Tutors in Taiwan. Teaching and Learning in Medicine, 17, (2), 179-183.

Lunt, I. , and Majors, K. (2000). The Professionalisation of Educational Psychology. Educational Psychology in Practice, 15, (4), 237-245.

Macdonald, R., and Savin-Baden, M. (2004). A Briefing on Assessment in Problem-based Learning. LTSN Generic Centre: Learning and Teaching Support Network.

Martenson, D., Myklebust, R. and Stalsburg, H. (1992). Implementing problem-based learning within a lecture-dominated curriculum: results and process. Teaching and Learning in Medicine, 4, 233-7.

Maudsley, G. (1999). Roles and responsibilities of the Problem-based learning tutor in the undergraduate medical curriculum. British Medical Journal, 318, 657-661.

Maudsley, G. (2002). Making Sense of Trying Not to teach: An Interview Study of Tutors' Ideas of Problem-based Learning. Academic Medicine, 77, (2), 162-172.

Mayo, W. P., Donnelly, M. B., Schwartz, R. W. (1995). Characteristics of the Ideal problem-based learning tutor in clinical medicine. Evaluation and the Health Professions, 18, (2), 124-136.

Mennin, S.P. and Kaufman, A. (1989). The change process and medical education. Medical Teacher, 11, 9-16.

Mennin,S. P., Friedman. M., Skipper, B., Kalishman, S., and Synder,J. (1993). Performances on the NBME I, II and III by medical students in the problem-based learning and conventional tracks at the University of New Mexico. Academic Medicine, 68, 616- 624.

Michaelsen, L. K., Fink, L. D., and Knight, A. (1997). Designing effective group activities: Lessons for classroom teaching and faculty development. In D. Dezure (ed), To improve the Academy (Vol 16, pp. 373-398).Fort Collins, CO: Professional and Organizational Development Network in Higher Education.

Monsen, J.J. (2000). Problem understanding within psychological interviewing. Unpublished PhD Thesis. Department of Psychology, University College London.

National Association of Educational Inspectors, Advisors and Consultants and The Association Of Educational Psychologists (NAEIAC / AEP) (2005). The Implications for Professionals of the DfES' Children's Workforce Strategy. Conference Report, April 2005.

Newman, M., Ambrose, K., Corner, T., Vernon, L., Quinn, S., Wallis, S., and Tymms, P. (2001). The Project on the Effectiveness of Problem Based Learning (PEPBL): A field trial in Continuing Professional Education. Paper presented at The Third International Inter-disciplinary Evidence-Based Policies and Indicator Systems Conference, July 2001.

Newman, M. (2004). A Pilot Systematic Review and Meta-Analysis on the Effectiveness of Problem Based Learning. School of Lifelong Learning and Education and School of Health and Social Sciences Middlesex University.

Norman, G. R. (1997). Assessments and problem based learning. In D. Boud, and G. Feletti (Eds.) The challenge of problem based learning, (2nd ed., pp. 262–268). London: Kogan Page.

Norman, G.R., Eva, K. W. and Schmidt, H.G. (2005). Implications of psychology type theories for full curriculum interventions. Medical Education, 39, (3), 247-249.

Norman, G.R., and Schmidt, H.G. (2000). Effectiveness of problem-based learning curricula: theory, practice and paper darts. Medical Education, 34, 721-728.

Premi, J., Shannon, S., Hartwick, K., Lamb, S., Wakefield, J. and Williams, J. (1994). Practice-based Small-group CME. Academic Medicine, 69, (10), 800-802.

Price, B. (2003). Studying nursing using problem-based and enquiry-based learning. Basingstoke: Palgrave McMillan.

Ramchandi, P., Joughlin, C., and Zwi, M. (2001). Evidence based child and adolescent mental health services: Oxymoron or brave new dawn. Child Psychology and Psychiatry Review, 6, 59-64.

Rudestam, K. E. and Newton, R.R. (1982). Surviving your Dissertation. Newbury Park.

Ryan, G. (1997). Ensuring that Students Develop an Adequate and Well-structured Knowledge Base. In D. Boud, and G. Feletti (Eds.), The challenge of problem based learning, (2nd ed., p. 125-136). London: Kogan Page.

Savin-Baden, M. (2000). Problem based learning in higher education: untold stories. Buckingham: The Society for Research into Higher Education / Open University Press.

Savin-Baden, M. (2003). Facilitating Problem-based Learning in Higher Education: Illuminating Perspectives. Buckingham: The Society for Research into Higher Education / Open University Press.

Savin-Baden, M. (2004). Understanding the impact of assessment on students in problem-based learning. Innovations in Education and Teaching International, 41, (2), 223-233.

Schmidt, H.G. (1994). Resolving inconsistencies in tutor expertise research: Lack of structure causes students to seek tutor guidance. Academic Medicine, 69, 656-662.

Schmidt, H.G., Van Der Arend, A., Moust, J.H.C., Kokx, I, and Boon, L. (1993). Influence of Tutor's Subject –matter Expertise on Student Effort and Achievement in Problem-based Learning. Academic Medicine, 68, (10), 784-791. (2000).

Schmidt, H. G, and Moust, J. H. C. (1995). What makes a Tutor Effective? A Structural-equations Modelling Approach to Learning in Problem-based Curricula. Academic Medicine, 70, (8), 708 - 714.

Schmidt, H. G. and Moust, J. H. C. (2000). Factors Affecting Small Group Tutorial Learning: A Review of Research. In Evensen, D. H. and Henlow, C. E. (2000). Problem based learning: A research perspective on learning interactions. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Schon, D. (1987). Educating the reflective practitioner: Towards a new design for teaching and learning in the professions. San Francisco: Jossey-Bass.

Schwartz, P.L., Heath,C. J., and Egan, A.,G. (1994). The Art of the Possible. Ideas from a traditional medical school engaged in curricular revision. University of Otago Press, Dunedin, New Zealand.

Schwartz, P., Mennin, S., and Webb, G. (Eds.) (2001). Problem based learning: Case studies, experience and practice. London: Kogan Page.

Schwartz, S. (2004). Time to bid goodbye to the psychology lecture. The Psychologist, 17, (1), 26-27.

Sherrard, C. (1997). Qualitative Research. The Psychologist, 10, 161-162.

Smits, P., Verbeek, J., and De Buisonje, C. (2002). Problem based learning in continuing medical education: a review of controlled evaluation studies. British Medical Journal, 324, 153-156.

Stefani, L., and Elton, L. (2002) Continuing professional development for academic teachers through self-initiated learning. Assessment and Evaluation in Higher Education, 27, (2), 117-129.

Stephenson, J., and Yorke. M. (1995). Capability and quality in higher education. London: Kogan- Page.

Stevenson, C., and Cooper., N. (1997). Qualitative and quantitative research. The Psychologist, 10, 159-160.

Strauss, A., and Corbin., J. (1990). Basis of Qualitative Research: Grounded Theory procedures and Techniques. London UK: Sage Publications.

Tait, H., and Entwistle, N. (1996). Identifying Students at risk through ineffective study strategies. Higher Education, 31, 97-116.

Thompson., D.G., and Williams., R.G. (1985). Barriers to the acceptance of problem-based learning in medical schools. Studies in Higher Education, 10, 199-204.

Toohey, S. (1999). Designing Courses for Higher Education. Buckingham: The Society for Research into Higher Education / Open University Press.

Tuckman, B. W., and Jensen, N. (1977). Stages of small group development revisited. Group and Organizational Studies, 2, 419-427.

Van Berkel, J.M., and Schmidt, H.G. (2000). Motivation to commit oneself as a determinant of achievement in problem-based learning. Higher Education, 40, 231-242.

Van den Bosshe, P., Gijbels, D., and Dochy, F. (2000). Does problem based learning educate problem solvers? A meta-analysis on the effects of problem based learning. Paper presented at the 7th EDINEB conference, Newport Beach USA June 2000.

Vernon, D. V. T. A. (1994). Tutor evaluation of Problem-based learning in Behavioural Sciences Courses. Teaching and Learning in Medicine, 6, 91-95.

Vernon, D. V. T. A. (1995). Attitudes and Opinions of faculty Tutors about problem-based Learning. Academic Medicine, 70, (3), 216-223.

Vernon, D.T. A., and Blake, R. L. (1993). Does problem based learning work? A meta-analysis of evaluative research. Academic Medicine, 69, (7), 550-563.

Vygotsky, L. (1978). Mind in society: The development of higher psychological processes. Cambridge MA: Harvard University Press.

Walton, H. J., and Matthews, M. B. (1989). Essentials of problem based learning. Medical Education, 23, 542-558.

Wetzel,M.,S. (1996). Developing the role of tutor/facilitator. Postgraduate Medical Journal, 72, 474-477.

Wilkerson, L., and Irby, D. M. (1998). Strategies for Improving Teaching Practice: A Comprehensive Approach to Faculty Development. Academic Medicine, 73, (4), 387-396.

Wilkerson, L., and Maxwell, J.A. (1988). A Qualitative Study of Initial Tutors In a Problem-Based Curriculum. Journal of Medical Education, 63, 892-899.

Wright, A. (1990). Educational psychologists and in-service training: Theory in Practice. In N. Jones and N. Frederickson (Eds.), Refocusing educational psychology (pp.104-129). Basingstoke: The Falmer Press.

Zinkiewicz, L., Hammond, N., and Trapp, A. (2003). Applying Psychology Disciplinary Knowledge to Psychology Teaching and Learning. LTSN Psychology. Report and Evaluation Series No 2, March 2003.